

; Entered [jdelaval 11-Feb-04 6:27]
SE01
acywkvct1

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 06:30:18 ; Search time 21 seconds
(without alignments)
36.636 Million cell updates/sec

Title: SEQ1

Perfect score: 54
Sequence: 1 acykwkvt 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	75.9	493	2	JC5486
2	40	74.1	322	2	AF1031
3	40	74.1	335	2	AI3453
4	40	74.1	475	2	T33943
5	40	74.1	3020	2	A43932
6	38	70.4	381	2	P59101
7	37	68.5	370	2	T47131
8	37	68.5	373	2	UC7289
9	37	68.5	1513	2	A54895
10	37	68.5	1567	2	T00568
11	36	66.7	12	1	UOGM2
12	36	66.7	12	1	S42765
13	36	66.7	12	2	US0423
14	36	66.7	12	2	US0424
15	36	66.7	13	2	PQ0445
16	36	66.7	83	2	S10706
17	36	66.7	125	2	I50499
18	36	66.7	125	2	I50498
19	36	66.7	253	2	T47303
20	36	66.7	627	2	B84482
21	36	66.7	765	2	T09574
22	36	66.7	1375	2	T37672
23	35	64.8	69	2	A86779
24	35	64.8	428	2	T48008
25	35	64.8	451	2	F83793
26	35	64.8	521	2	T38466
27	35	64.8	871	2	S68482
28	35	64.8	940	2	H86420
29	35	64.8	1360	2	T12064

30	34.5	63.9	235	2	T44466
31	34	63.0	104	2	S69756
32	34	63.0	392	2	T46418
33	34	63.0	421	2	T22969
34	34	63.0	504	2	A38221
35	34	63.0	506	2	A40679
36	34	63.0	523	2	B40679
37	34	63.0	654	2	AH2445
38	34	63.0	2704	2	S08118
39	34	63.0	2718	2	A23475
40	33.5	62.0	187	2	T10073
41	33.5	62.0	258	2	T30368
42	33	61.1	87	2	T33285
43	33	61.1	100	2	A31201
44	33	61.1	140	2	T33284
45	33	61.1	155	2	A99373

ALIGNMENTS

RESULT 1
JC5486
membrane glycoprotein CD30 homolog precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 07-Dec-1999
C:Accession: JC5486
R:Alizawa, S.; Satoh, H.; Horie, R.; Ito, K.; Choi, S.H.; Takeuchi, H.; Watanabe, T.
Gene 182, 155-162, 1996
A:Title: Cloning and characterization of a cDNA for rat CD30 homolog and chromosomal as-
A:Reference number: JC5486; MUID:97136705; PMID:8982082
A:Contents: T cell
A:Accession: JC5486
A:Molecule type: mRNA
A:Residues: 1-493 <AI2>
F:1-16/Domain: signal sequence #status predicted <SIG>
F:19-493/Product: membrane glycoprotein CD30 homolog #status predicted <TM>
F:278-304/Domain: transmembrane #status predicted <TM>
F:305-493/Domain: intracellular #status predicted <INT>

Query Match 75.9%; Score 41; DB 2; Length 493;
Best local similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWVC 7
DB 304 CYWKAC 309

RESULT 2
AF1031
hypothetical protein SRY4576 (imported) - Salmonella enterica subsp. enterica serovar Ty-
C:Species: Salmonella enterica subsp. enterica serovar Typh-
A:Note: this species has also been called Salmonella typh-
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AF1031
R:Fairhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov-
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AF1031
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <PAR>
A:Cross-references: GB:AL51382; PDB:CAD09351.1; PID:g16505351; GSPDB:GN00176
C:Genetics:
A:Gene: SRY4576

Query Match 74.1%; Score 40; DB 2; Length 322;

Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWVCT 8
|||:
DB 49 CYWVCT 55

RESULT 3

transcription regulator, arac family BME1615 (imported) - Brucella melitensis (strain 1
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #ext_change 01-Feb-2002

C/Accession: A13453
R/DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Loe, T.; Ivanova,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A/Reference number: AD3252; PMID:11756688
A/Accession: A13453

A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-335 <RUR>

A/Cross-references: GB:AE008917; PIDN:AAU52796.1; PID:G17983633; GSPDB:GN00190
A/Experimental source: strain 16M

C/Genetics:

A/Gene: BME1615
A/Map position: 1

Query Match 74.1%; Score 40; DB 2; Length 335;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWVCT 8
|||:
DB 60 CAWKICT 66

RESULT 4

hypothetical protein C01B4.7 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #ext_change 29-Oct-1999

C/Accession: T33943
R/Smith, A.; Mameley, P.; Fromick, W.

Submitted to the EMBL Data Library, February 1999
A/Description: The sequence of C. elegans cosmid C01B4.

A/Reference number: Z21443

A/Accession: T33943

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-475 <SMI>

A/Cross-references: EMBL:AF129952; PIDN:AAU14699.1; GSPDB:GN00023; CESP:C01B4.7

A/Experimental source: strain Bristol N2; clone C01B4

C/Genetics:

A/Gene: CESP:C01B4.7

A/Map position: 5

A/Introns: 45/2; 80/1; 118/2; 189/3; 239/2; 340/3; 433/3

Query Match 74.1%; Score 40; DB 2; Length 475;
Best Local Similarity 83.3%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWVCT 8
|||:
DB 264 YWVCT 269

RESULT 5

muslin 2 precursor, intestinal - human (fragments)

N/Alternate names: muslin SMUC-41

C/Species: Homo sapiens (man)

C/Date: 10-Mar-1993 #sequence_revision 12-Apr-1996 #ext_change 05-Nov-1999
C/Accession: A49963; A45106; B45106; A43932; B3532; A61257; PQ0328; PQ0329
R/Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Siddiki, B.; Kim, Y.S.

J. Biol. Chem. 269, 2440-2446, 1994

A/Title: Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the

A/Reference number: A49963; PMID:94132002; PMID:8300571

A/Accession: A49963

A/Molecule type: mRNA

A/Residues: 1-639 <GU1>

A/Cross-references: GB:L21998

R/Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Roche, E.M.; Lagace, R.E.; Kim, Y.S.

J. Biol. Chem. 267, 21375-21383, 1992

A/Title: The human MUC2 intestinal mucin has cysteine-rich subdomains located both upstr

A/Reference number: A45106; PMID:93016075; PMID:1400449

A/Accession: A45106

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 626-1895 <GU2>

A/Cross-references: GB:M94131; NID:G186395; PIDN:AAAS9163.1; PID:G186396

A/Note: sequence extracted from NCBI backbone (NCBIP:116706)

A/Accession: B45106

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 2037-3020 <GU3>

A/Cross-references: GB:M94132; NID:G186397; PIDN:AAAS9164.1; PID:G186398

A/Experimental source: colon

A/Note: sequence extracted from NCBI backbone (NCBIP:116698)

R/Toribara, N.W.; Gum Jr., J.R.; Culhane, P.J.; Lagace, R.E.; Hicks, J.W.; Petersen, G.M

J. Clin. Invest. 89, 1005-1013, 1991

A/Title: MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymorp

A/Reference number: A43932; PMID:91358717; PMID:1885763

A/Accession: A43932

A/Molecule type: DNA

A/Residues: 1343-1350, 'L', 1352-1411, 'S', 1413-1448, 'P', 1450-1503, 'T', 1505-1915 <TOR>

A/Cross-references: GB:M74027; NID:G186863; PIDN:AAAS9875.1; PID:G186864

A/Note: sequence inconsistent with the nucleotide translation

A/Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)

R/Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampert, D.T.A.; Kim, Y.S.

J. Biol. Chem. 264, 6480-6487, 1989

A/Title: Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and eviden

A/Reference number: A3532; PMID:89197956; PMID:2703501

A/Accession: B3532

A/Molecule type: mRNA

A/Residues: 1916-2193 <GU4>

A/Cross-references: GB:M2405; NID:G186873; PIDN:AAA36334.1; PID:G186874

A/Experimental source: intestine

R/Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.; Basbaum, C.B.

J. Clin. Invest. 87, 77-82, 1991

A/Title: Human bronchus and intestine express the same mucin gene.

A/Reference number: A61257; PMID:91086481; PMID:1985113

A/Accession: A61257

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 'T', 1925-1948, 'TTS', 1952-1954 <JAN>

A/Experimental source: bronchus

R/Xu, G.; Huan, L.; Khatri, I.; Sajjan, U.S.; McCool, D.; Wang, D.; Jones, C.; Foretner,

Biochem. Biophys. Res. Commun. 183, 821-826, 1992

A/Title: Human intestinal mucin-like protein (MLP) is homologous with rat MLP in the C-t

A/Reference number: PQ0328; PMID:92198477; PMID:1550588

A/Accession: PQ0328

A/Molecule type: mRNA

A/Residues: 2328-2468 <XUG>

A/Cross-references: GB:M86523

A/Experimental source: small intestine

A/Accession: PQ0329

A/Molecule type: protein

A/Residues: 2328-2342, 'K', 2344-2354 <XUG1>

C/Genetics:

A/Gene: GDB:MUC2

A/Cross-references: GDB:120203; OMIM:158370

A/Map position: 11p15.5-11p15.5

C/Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; von

C/Keywords: glycoprotein; intestine; tandem repeat

F:2766-2834/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 74.1%; Score 40; DB 2; Length 3020;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYMKVC 7
DB 1238 CYMECT 1243

RESULT 6

hypothetical protein pXOI-86 - Bacillus anthracis virulence plasmid pXOI
C:Species: Bacillus anthracis
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #ext_change 11-May-2000
C:Accession: F59101
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
J. Bacteriol. 181, 6509-6515, 1999
A:Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbored
A:Reference number: A59091; MUID:99445483; PMID:10515943
A:Accession: F59101
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-381 <OKI>
A:Cross-references: GB:A065404; NID:g4894216; PIDN:AD32390.1; PID:g4894302
A:Experimental source: strain Sterne
A>Note: Similar to hypothetical protein in BLTR-SpoIIIC intergenic region, yrKO B. subtilis
C:Genetics:
A:Gene: pXOI-86
A:Genome: plasmid

Query Match 70.4%; Score 38; DB 2; Length 381;
Best Local Similarity 63.3%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YMKVCT 8
DB 340 YMKLCT 345

RESULT 7

G-protein coupled receptor, SREB2 - human
C:Species: Homo sapiens (man)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #ext_change 21-Jun-2002
C:Accession: T47131; JC7288
R:Pouletka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24374
A:Accession: T47131
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-370 <AAA>

A:Cross-references: EMBL:AL161959; NID:g7328012; PIDN:CA882307.1; PID:g7328013
A:Experimental source: adult amygdala; clone DKFZp761L08121
R:Matsumoto, M.; Saito, T.; Takasaki, J.; Kamohara, M.; Sugimoto, T.; Kobayashi, M.; Tad
Biochem. Biophys. Res. Commun. 272, 576-582, 2000
A:Title: An evolutionarily conserved G-protein coupled receptor family, SREB, expressed
A:Reference number: JC7287
A:Accession: JC7288
A:Molecule type: mRNA
A:Residues: 1-370 <MAT>
A:Cross-references: DDBJ:AB040799

C:Genetics:
A:Gene: sreB2
A:Map position: 7q31
A>Note: DKFZp761L08121.1

C:Keywords: brain; G protein-coupled receptor; glycolysis; reproduction; transmembrane p
Query Match 68.5%; Score 37; DB 2; Length 370;
Best Local Similarity 83.3%; Pred. No. 81;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACYMKV 6
DB 304 ACYMKV 309

RESULT 8

G-protein coupled receptor, SREB3 - human
C:Species: Homo sapiens (man)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #ext_change 19-Apr-2002
C:Accession: JC7289
R:Matsumoto, M.; Saito, T.; Takasaki, J.; Kamohara, M.; Sugimoto, T.; Kobayashi, M.; Tad
Biochem. Biophys. Res. Commun. 272, 576-582, 2000
A:Title: An evolutionarily conserved G-protein coupled receptor family, SREB, expressed
A:Reference number: JC7287
A:Accession: JC7289
A:Molecule type: mRNA
A:Residues: 1-373 <MAT>
A:Cross-references: DDBJ:AB040799
C:Genetics:
A:Gene: sreB3
A:Map position: Xp11
C:Superfamily: endochelin receptor B
C:Keywords: brain; G protein-coupled receptor; glycolysis; reproduction; transmembrane p

Query Match 68.5%; Score 37; DB 2; Length 373;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACYMKV 6
DB 305 ACYMKV 310

RESULT 9

mucin 2, intestinal/tracheal - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #ext_change 11-Jan-2000
C:Accession: A54895
R:Ohmori, H.; Dohman, A.F.; Gallup, M.; Tsuda, T.; Kai, H.; Gum Jr., J.R.; Kim, Y.S.; I
J. Biol. Chem. 269, 17833-17840, 1994
A:Title: Molecular cloning of the amino-terminal region of a rat MUC 2 mucin gene homolo
A:Reference number: A54895; MUID:94299489; PMID:8027037
A:Accession: A54895
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1513 <OHM>
A:Cross-references: GB:U07615
C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; von
C:Keywords: intestine

Query Match 68.5%; Score 37; DB 2; Length 1513;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
DB 1235 CYMECT 1240

RESULT 10

hypothetical protein At2g39580 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F12L6.24
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #ext_change 16-Feb-2001
C:Accession: T00568; B84819
R:Bounslay, S.D.; Jin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kau
submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC F12L6 genomic sequence.
A:Reference number: Z14168

A:Accession: T00568
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1567 <ROU>
 A:Cross-references: EMBL:AC004218; NID:g3355463; PID:g3355487
 A:Experimental source: Cultivar Columbia
 A:Rutin, X.; Kaul, S.; Rounbley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
 Neus, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: B84819
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1567 <STO>
 A:Cross-references: GB:AE002093; NID:g3355487; PID:AA027849.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g39580; F12L6.24
 A:Map position: 2
 A:Introns: 306/1; 346/3; 395/3; 688/1; 843/1; 945/3; 991/3; 1098/1
 C:Superfamily: Arabidopsis thaliana hypothetical protein F12L6.24

Query Match
 Best Local Similarity 68.5%; Score 37; DB 2; Length 1567;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 |||
 Db 1539 CYWSLC 1544

RESULT 11
 UOGEN2
 urotensin II - long-jawed mudsucker
 C:Species: Gilllichthys mirabilis (long-jawed mudsucker)
 C:Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 15-Oct-1996
 C:Accession: A01409
 R:Pearson, D.; Shively, J.E.; Clark, B.R.; Geschwind, I.I.; Barkley, M.; Nishio, R.; E
 Proc. Natl. Acad. Sci. U.S.A. 77: 5021-5024, 1980
 A:Title: Urotensin II: a somatostatin-like peptide in the caudal neurosecretory system c
 A:Reference number: A01409; MUID:81054904; PMID:6107911
 A:Accession: A01409
 A:Molecule type: protein
 A:Residues: 1-12 <PRA>
 A:Note: the proposed sequence was confirmed by synthesis of a peptide with the same stru
 C:Comment: Urotensin II is found in the teleost caudal neurosecretory system and is inv
 C:Superfamily: urotensin II
 C:Keywords: neuropeptide; osmoregulation
 F:6-11/Disulfide bonds: #status experimental

Query Match
 Best Local Similarity 66.7%; Score 36; DB 1; Length 12;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 |||
 Db 6 CFMKVC 11

RESULT 12
 S42765
 urotensin II - teleostean fish
 C:Species: teleostean fish
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Dec-1997
 C:Accession: S42765
 R:Bhaakaran, R.; Arunkumar, A.I.; Yu, C.
 Biochim. Biophys. Acta 1199, 115-122, 1994
 A:Title: NMR and dynamical simulated annealing studies on the solution conformation of u
 A:Reference number: S42765; MUID:94169160; PMID:8123660
 A:Accession: S42765
 A:Molecule type: protein
 A:Residues: 1-12 <BHA>

C:Superfamily: urotensin II
 C:Keywords: neuropeptide; osmoregulation
 F:6-11/Product: urotensin II #status experimental <MAT>

Query Match
 Best Local Similarity 66.7%; Score 36; DB 2; Length 12;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 |||
 Db 6 CFMKVC 11

RESULT 13
 JS0423
 urotensin II-A peptide - white sucker
 C:Species: Catostomus commersoni (white sucker)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-May-1997
 C:Accession: JS0423
 R:McMaster, D.; Lederis, K.
 Peptides 4, 367-373, 1983
 A:Title: Isolation and amino acid sequence of two urotensin II peptides from Catostomus
 A:Reference number: JS0423; MUID:84041959; PMID:6138758
 A:Accession: JS0423
 A:Molecule type: protein
 A:Residues: 1-12 <MCM>
 C:Comment: This peptide has smooth muscle-stimulating activity.
 C:Superfamily: urotensin II
 F:6-11/Disulfide bonds: #status experimental

Query Match
 Best Local Similarity 66.7%; Score 36; DB 2; Length 12;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 |||
 Db 6 CFMKVC 11

RESULT 14
 JS0424
 urotensin II-B peptide - white sucker
 C:Species: Catostomus commersoni (white sucker)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
 C:Accession: JS0424
 R:McMaster, D.; Lederis, K.
 Peptides 4, 367-373, 1983
 A:Title: Isolation and amino acid sequence of two urotensin II peptides from Catostomus
 A:Reference number: JS0423; MUID:84041959; PMID:6138758
 A:Accession: JS0424
 A:Molecule type: protein
 A:Residues: 1-12 <MCM>
 C:Comment: This peptide has smooth muscle-stimulating activity.
 F:6-11/Disulfide bonds: #status experimental

Query Match
 Best Local Similarity 66.7%; Score 36; DB 2; Length 12;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 |||
 Db 6 CFMKVC 11

RESULT 15
 PQ0445
 urotensin II - laughing frog
 C:Species: Rana ridibunda (laughing frog)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 03-Mar-1995
 C:Accession: PQ0445
 R:Conlon, J.M.; O'Harte, F.; Smith, D.D.; Tonon, M.C.; Vaudry, H.
 Biochem. Biophys. Res. Commun. 188, 576-583, 1992
 A:Title: Isolation and primary structure of urotensin II from the brain of a tetrapod, r

A:Reference number: PQ0445; MUID:93075134; PMID:1445302

A:Accession: PQ0445

A:Molecule type: protein

A:Residues: 1-13 <CON>

A:Experimental source: brain

C:Superfamily: urotensin II

Query Match

66.7%; Score 36; DB 2; Length 13;

Best Local Similarity 66.7%; Pred. No. 9;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CYMKVC 7

Db 7 CFMKYC 12

Search completed: February 11, 2004, 06:30:56
Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 06:30:18 ; Search time 11 Seconds

(without alignments)
34.201 Million cell updates/sec

Title: SEQ1

Perfect score: 54

Sequence: 1 acykwvct 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	75.9	493	1	TNR8_RAT
2	41	75.9	498	1	TNR8_MOUSE
3	40	74.1	5179	1	MUC2_HUMAN
4	39	72.2	387	1	SRB3_BRARE
5	37	68.5	332	1	MCAR_PIG
6	37	68.5	370	1	GPB5_HUMAN
7	37	68.5	371	1	GPB5_BRARE
8	37	68.5	373	1	SRB3_HUMAN
9	37	68.5	373	1	SRB3_RAT
10	37	68.5	520	1	VNN2_HUMAN
11	37	68.5	1513	1	MUC2_RAT
12	36	66.7	12	1	UR2A_CATCO
13	36	66.7	12	1	UR2B_CATCO
14	36	66.7	12	1	UR2B_CYPCA
15	36	66.7	12	1	UR2_GILMI
16	36	66.7	12	1	UR2_POLSP
17	36	66.7	12	1	UR2_SCYCA
18	36	66.7	29	1	CYOC_VIOOD
19	36	66.7	29	1	KAB5_OLDAF
20	36	66.7	29	1	VARA_VIOAR
21	36	66.7	29	1	VARC_VIOAR
22	36	66.7	29	1	VARD_VIOAR
23	36	66.7	29	1	VARE_VIOAR
24	36	66.7	83	1	UR2_PLAPE
25	36	66.7	123	1	UR2_MOUSE
26	36	66.7	123	1	UR2_RAT
27	36	66.7	124	1	UR2_HUMAN
28	36	66.7	125	1	UR2A_CYPCA
29	36	66.7	125	1	UR2G_CYPCA
30	36	66.7	127	1	UR2_RANRI
31	36	66.7	765	1	S23A_HUMAN
32	36	66.7	765	1	S23A_MOUSE
33	35	64.8	124	1	KAB1_OLDAF

34	35	64.8	166	1	VSN1_NOCAR	P50186 nocardia ae
35	35	64.8	210	1	KAB2_OLDAF	P58454 oldenlandia
36	35	64.8	466	1	ZIC3_MOUSE	O62521 mus musculu
37	35	64.8	467	1	ZIC3_HUMAN	O60481 homo sapien
38	35	64.8	521	1	YAV8_HUMAN	Q10177 schizosach
39	35	64.8	871	1	SC10_YEAST	O06245 saccharomyc
40	35	64.8	1360	1	GLI1_XENLA	Q91690 xenopus lae
41	34	63.0	450	1	FEM2_RAT	O99w77 rattus norv
42	34	63.0	504	1	CHT_BRUMA	P29030 brugia mala
43	34	63.0	663	1	GLI3_CHICK	P55879 gallus gall
44	34	63.0	1258	1	GLI2_HUMAN	P10070 homo sapien
45	34	63.0	1596	1	GLI3_MOUSE	O61602 mus musculu

ALIGNMENTS

RESULT 1	ID	TNR8_RAT	STANDARD	PRT	493 AA.
AC	P97525	TNR8_RAT			
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Tumor necrosis factor receptor superfamily member 8 precursor (CD30L receptor) (lymphocyte activation antigen CD30).				
GN	TNFRSF8 OR CD30.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_Taxid=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAINE=MKAH; TISSUE=T-cell lymphoma;				
RX	MEDLINE=97136705; PubMed=8982082;				
RA	Aizawa S., Sato H., Horie R., Ito K., Choi S.H., Takeuchi H.,				
RA	Watanabe T.;				
RT	*Cloning and characterization of a cDNA for rat CD30 homolog and				
RT	chromosomal assignment of the genomic gene.";				
RL	Gene 182:155-162(1996).				
CC	-1- FUNCTION: Receptor for TNFRSF8/CD30L. May play a role in				
CC	the regulation of cellular growth and transformation of				
CC	activated lymphoblasts. Regulates gene expression through				
CC	activation of NF-kappa-B (By similarity).				
CC	-1- SUBUNIT: Interacts with TRAF1, TRAF2, TRAF3 and TRAF5 (By				
CC	similarity).				
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).				
CC	-1- TISSUE SPECIFICITY: Very low level of expression. Detected in				
CC	spleen, thymus and lung. Highly expressed in HTLV-1 infected T-				
CC	cell lines.				
CC	-1- INDUCTION: By phytohemagglutinin (PHA) in spleen T-cells.				
CC	-1- SIMILARITY: Contains 3 TNFR-Cys repeats.				
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL; DA2117; BA007699.1; -				
DR	PIR; JC5486; JC5486.				
DR	Interpro; IPR001368; TNFR_c6.				
DR	Pfam; PF00020; TNFR_c6; 3.				
DR	SMART; SM00208; TNFR_2.				
DR	PROSITE; PS00652; TNFR_NGFR_1; FALSE_NEG.				
DR	PROSITE; PS00650; TNFR_NGFR_2; 1.				
KW	Receptor; Transmembrane; Glycoprotein; Signal; Repeat.				
FT	CHAIN	1	18		POTENTIAL.
FT	FT	1	493		TUMOR NECROSIS FACTOR RECEPTOR
FT	FT	19	255		SUPERFAMILY MEMBER 8.
FT	FT				EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 256 276 POTENTIAL.
FT DOMAIN 277 493 CYTOPLASMIC (POTENTIAL).
FT REPEAT 28 66 TNFR-CYS 1.
FT REPEAT 68 106 TNFR-CYS 2.
FT REPEAT 107 150 TNFR-CYS 3.
FT DISULFID 29 44 BY SIMILARITY.
FT DISULFID 45 58 BY SIMILARITY.
FT DISULFID 84 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 107 121 BY SIMILARITY.
FT DISULFID 124 142 POTENTIAL.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 493 AA; 52658 MW; 408B7503B3CB87C CRC64;

Query Match 75.9%; Score 41; DB 1; Length 493;
Best Local Similarity 83.3%; Pred. No. 6.7;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CYMKVC 7
Db 304 CYMKAC 309

```

```

RESULT 2
TNFR_MOUSE
ID TNFR_MOUSE STANDARD; PRT; 498 AA.
AC 060846;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 8 precursor (CD30L
DE receptor) (lymphocyte activation antigen CD30).
GN TNFRSF8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Splenic;
RA MEDLINE=96132967; PubMed=8543792;
RX Bowen M.A., Lee R.K., Miragiotou G., Nam S.Y., Podack E.R.;
RT "Structure and expression of murine CD30 and its role in cytokine
RT production.";
RL J. Immunol. 156:442-449 (1996).
CC -1- FUNCTION: Receptor for TNFRSF8/CD30L. May play a role in
CC the regulation of cellular growth and transformation of
CC activated lymphoblasts. Regulates gene expression through
CC activation of NF-kappa-B (By similarity).
CC -1- SUBUNIT: Interacts with TRAF1, TRAF2, TRAF3 and TRAF5 (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- TISSUE SPECIFICITY: Detected in thymus and in activated
CC splenocytes.
CC -1- INDUCTION: By concanavalin A and pokeweed mitogen in splenocytes.
CC -1- SIMILARITY: Contains 3 TNFR-Cys repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL, U25416; AAA92887.1; -
CC MGD: MGI:99908; Tnfref8.
CC InterPro: IPR001368; TNFR_c6.
CC Pfam: PF00020; TNFR_c6; 3.
CC SMART; SM00208; TNFR; 3.
CC PROSITE; PS00652; TNFR_NGFR_1; 1.

```

```

DR PROSITE; PS00650; TNFR_NGFR_2; 2.
KM Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 498 TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 8.
FT DOMAIN 19 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 279 POTENTIAL.
FT DOMAIN 280 498 CYTOPLASMIC (POTENTIAL).
FT REPEAT 28 66 TNFR-CYS 1.
FT REPEAT 68 106 TNFR-CYS 2.
FT REPEAT 107 150 TNFR-CYS 3.
FT DISULFID 29 44 BY SIMILARITY.
FT DISULFID 45 58 BY SIMILARITY.
FT DISULFID 48 65 BY SIMILARITY.
FT DISULFID 69 81 BY SIMILARITY.
FT DISULFID 84 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 107 121 BY SIMILARITY.
FT DISULFID 128 146 BY SIMILARITY.
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 498 AA; 53216 MW; 98CA2A05B38AFA71 CRC64;

Query Match 75.9%; Score 41; DB 1; Length 498;
Best Local Similarity 83.3%; Pred. No. 6.8;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CYMKVC 7
Db 308 CYMKAC 313

```

```

RESULT 3
MUC2_HUMAN
ID MUC2_HUMAN STANDARD; PRT; 5179 AA.
AC 002817; Q14878;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mucin 2 precursor (Intestinal mucin 2).
GN MUC2 OR SMUC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=94132002; PubMed=8300571;
RA Gum J.R. Jr., Hicks J.W., Toribara N.W., Siddiki B., Kim Y.S.;
RT "Molecular cloning of human intestinal mucin (MUC2) cDNA.
RT Identification of the amino terminus and overall sequence similarity
RT to prepro-von Willebrand factor.";
RL J. Biol. Chem. 269:2440-2446 (1994).
RN [2]
RP SEQUENCE OF 626-1895 AND 4196-5179 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=93016075; PubMed=1400449;
RA Gum J.R. Jr., Hicks J.W., Toribara N.W., Roche E.-M., Lagace R.E.,
RA Kim Y.S.;
RT "The human MUC2 intestinal mucin has cysteine-rich subdomains located
RT both upstream and downstream of its central repetitive region.";
RL J. Biol. Chem. 267:21375-21383 (1992).
RN [3]
RP SEQUENCE OF 1343-1895 AND 4176-4195 FROM N.A.
RX MEDLINE=91358717; PubMed=1885763;
RA Toribara N.W., Gum J.R. Jr., Culhane P.J., Lagace R.E., Hicks J.W.,
RA Petersen G.M., Kim Y.S.;
RT "MUC-2 human small intestinal mucin gene structure. Repeated arrays
RT and polymorphism.";
RL J. Clin. Invest. 88:1005-1013 (1991).
CC -1- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND

```

```

CC OTHER MUCUS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS
CC AGENTS AT MUCOSAL SURFACES.
CC -1- SUBUNIT: MULTIMERIC.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,
CC BRONCHUS, CERVIX AND GALL BLADDER.
CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -1- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND
CC VARIES AMONG DIFFERENT ALLELES.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -1- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
CC -1- SIMILARITY: Contains 2 VWFC domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, I21998; AAB95295.1; -
CC EMBL, M74027; AAA59875.1; -
CC EMBL, M94131; AAA59164.1; -
CC EMBL, M94132; AAA59164.1; -
CC PIR, A49963; A43932.
CC Genew; HGNC:7512; MUC2.
CC MIM, 158370; -
CC GO: 0005803; C:secretory vesicle; TAS.
CC InterPro: IPR006208; Cys_knot.
CC InterPro: IPR006207; Cys_knot_C.
CC InterPro: IPR006209; EGF_like_C.
CC InterPro: IPR002919; TIL_Cystich.
CC InterPro: IPR001007; VWF_D.
CC InterPro: IPR001846; VWF_D.
CC Pfam; PF00007; Cys_knot; 1.
CC Pfam; PR01826; TIL; 1.
CC Pfam; PF00093; vwc; 1.
CC Pfam; PF00094; vwc; 4.
CC SMART; SM00214; vwc; 2.
CC SMART; SM00216; vwc; 4.
CC PROSITE; PS01185; CTCK_1; 1.
CC PROSITE; PS00022; EGF_1; UNKNOWN_1.
CC PROSITE; PS01208; VWFC_1; 2.
CC PROSITE; PS0184; VWFC_2; 2.
CC Glycoprotein; Repeat; Signal.
CC FT SIGNAL 1 20 POTENTIAL.
CC FT CHAIN 21 5179 MUCIN 2.
CC FT DOMAIN 1401 1747 APPROXIMATE REPEATS.
CC REPEAT 1401 1416 1.
CC REPEAT 1417 1432 2.
CC REPEAT 1433 1448 3.
CC REPEAT 1449 1464 4.
CC REPEAT 1465 1471 5.
CC REPEAT 1472 1478 6.
CC REPEAT 1479 1494 7A.
CC REPEAT 1495 1517 7B.
CC REPEAT 1518 1533 8A.
CC REPEAT 1534 1556 8B.
CC REPEAT 1557 1572 9A.
CC REPEAT 1573 1596 9B.
CC REPEAT 1597 1612 10A.
CC REPEAT 1613 1635 10B.
CC REPEAT 1636 1651 11A.
CC REPEAT 1652 1675 11B.
CC REPEAT 1676 1693 12.
CC REPEAT 1694 1699 13.
CC REPEAT 1700 1715 14.
CC REPEAT 1716 1731 15.
CC REPEAT 1732 1747 16.

```

```

FT DOMAIN 4815 4886 'VWFC 1.
FT DOMAIN 4924 4991 'VWFC 2.
FT DISULFID 5075 5160 CTCK.
FT DISULFID 5075 5122 BY SIMILARITY.
FT DISULFID 5089 5136 BY SIMILARITY.
FT DISULFID 5098 5152 BY SIMILARITY.
FT DISULFID 5102 5154 BY SIMILARITY.
FT DISULFID 5159 5159 BY SIMILARITY.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 770 770 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 894 894 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1154 1154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1215 1215 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1230 1230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1246 1246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1787 1787 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1820 1820 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4339 4339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4351 4351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4362 4362 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4373 4373 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4422 4422 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4438 4438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4502 4502 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4616 4616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4627 4627 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4752 4752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4787 4787 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4881 4881 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4888 4888 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4955 4955 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4970 4970 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5019 5019 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5038 5038 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5069 5069 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1351 1351 H -> L (IN REF. 3).
FT CONFLICT 1412 1412 T -> S (IN REF. 3).
FT CONFLICT 1449 1449 L -> P (IN REF. 3).
FT CONFLICT 1504 1504 M -> T (IN REF. 3).
FT CONFLICT 4192 4192 G -> S (IN REF. 2).
SQ SEQUENCE 5179 AA; 540295 MW; 85CD571FB9A5653 CRC64;

Query Match 74.1%; Score 40; DB 1; Length 5179;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYMKVC 7
Db 1238 CYWEIC 1243

RESULT 4
SRB3_BRARE STANDARD; PRT; 387 AA.
AC Q91918;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Super conserved receptor expressed in brain 3.
GN SRB3.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20294882; PubMed=10833454;
RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,
RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuchi K.;

```

Qy	Db	1 ACYKWTCT 8	171 SCTWAVCT 178
68.5%	Score 37;	DB 1;	Length 332;
62.5%	Pred. No. 22;		
5;	Conservative	1;	Mismatches
2;	Indels	0;	Gaps
0;			

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable G protein-coupled receptor GPR85 (Super conserved receptor
 DE expressed in brain 2) (PRTX1).
 GN GPR85 OR SREB2.
 OS Homo sapiens (Human),
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606, 10090, 10116;
 (1)
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human, and Mouse; STRAIN=CD-1; TISSUE=Brain, and Fetal brain;
 RX MEDLINE=20435311; PubMed=10978537;
 RA Hellebrand S., Schaller H.C., Wittenberger T.;
 RT "The brain-specific G-protein coupled receptor GPR85 with identical
 RL protein sequence in man and mouse maps to human chromosome 7q31.";
 RL Biochim. Biophys. Acta 1493:269-272(2000).
 (2)
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human, and Rat; TISSUE=Brain;
 RX MEDLINE=20294882; PubMed=10833454;
 RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,
 RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuchi K.;
 RT "An evolutionarily conserved G-protein coupled receptor family, SREB,
 RL expressed in the central nervous system.";
 RL Biochem. Biophys. Res. Commun. 272:576-582(2000).
 (3)
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RX Suwa M., Sato T., Okouchi I., Arita M., Putani K., Matsumoto S.,
 RA Tautumi S., Aburatani H., Asai K., Akiyama Y.;
 RT "Genome-wide discovery and analysis of human seven transmembrane helix
 RL receptor genes.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 (4)
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human; TISSUE=Amygdala;
 RX Pouskka A., Wellenreuther R., Newes H.-W., Weil B., Wismann S.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 (5)
 RP SEQUENCE FROM N.A.
 RC SPECIES=Rat; STRAIN=Sprague-Dawley;
 RX Kim H., Park S., Kang Y., Kim C., Jeon J.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 (6)
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=C57BL/6; TISSUE=Eye;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heib F.,
 RA Diatchenko L., Mausina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalski U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: ORPHAN RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- TISSUE SPECIFICITY: IN HUMAN HIGHLY EXPRESSED IN BRAIN AND TESTIS.

CC LOWER LEVELS IN SMALL INTESTINE, PLACENTA AND SPLEEN. IN BRAIN
 CC REGIONS, DETECTED IN ALL REGIONS TESTED, BUT SOMEWHAT LOWER LEVELS
 CC IN THE CORPUS CALLOSUM, MEDULLA AND SPINAL CORD. IN THE MOUSE
 CC EXCLUSIVELY EXPRESSED IN BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF250237; AAF79956.1; -
 CC EMBL: AB040800; BAA96646.1; -
 CC EMBL: AL161959; CAB82307.1; -
 CC EMBL: AF254416; AAF79959.1; -
 CC EMBL: AB065688; BAC05911.1; -
 CC EMBL: AB040803; BAA96649.1; -
 CC EMBL: AF203907; AAG42284.1; -
 CC EMBL: BC026975; AAB26975.1; -
 CC EMBL: T47131; T47131.
 CC PIR: T47131; T47131.
 CC Genew: HGNC:4536; GPR85.
 CC MIM: 605188; -
 CC MGD: MGI:1927651; Gpr85.
 CC GO: GO:0007165; P:signal transduction; TMS.
 CC Interpro: IPR000276; GPCR_Rhodopsn.
 CC Pfam: PF00001; 7tm_1, 1.
 CC PRINTS: PR00237; GPCR_Rhodopsn.
 CC PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; FALSE_NEG.
 CC PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Multigene family.
 CC FT DOMAIN 1 25
 CC FT TRANSMEM 26 46
 CC FT DOMAIN 47 57
 CC FT TRANSMEM 58 78
 CC FT DOMAIN 79 96
 CC FT TRANSMEM 97 117
 CC FT DOMAIN 118 137
 CC FT TRANSMEM 138 158
 CC FT DOMAIN 159 188
 CC FT TRANSMEM 189 209
 CC FT DOMAIN 210 286
 CC FT TRANSMEM 287 307
 CC FT DOMAIN 308 313
 CC FT TRANSMEM 314 334
 CC FT DOMAIN 335 370
 CC FT DISULFID 94 172
 CC FT CARBOHYD 3 3
 CC FT CARBOHYD 83 83
 CC FT CARBOHYD 182 182
 CC SEQUENCE 370 AA; 41994 MW; 7B67A39F6166AAB CRC64;
 CC
 CC Query Match 68.5%; Score 37; DB 1; Length 370;
 CC Best Local Similarity 83.3%; Pred. No. 24;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 ACYMKV 6
 CC DB 304 ACYMKV 309
 CC
 CC RESULT 7
 CC GP85_BRARE STANDARD; PRT; 371 AA.
 CC ID 091919;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Super conserved receptor expressed in brain 2.
 CC GN SREB2.

OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20294882; PubMed=10833454;
 RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,
 RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuchi K.,
 RT "An evolutionarily conserved G-protein coupled receptor family, SRBB,
 RT expressed in the central nervous system.";
 RL Biochem. Biophys. Res. Commun. 272:576-582 (2000).
 CC
 CC -1- FUNCTION: ORPHAN RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL, AB040805; BAA96651.1; -;
 CC ZFIN, ZDB-GENE-000710-2; strb2.
 CC InterPro, IPR000276; GPCR_Rhodopsin.
 DR Pfam, PF00001; 7tm_1, 1.
 DR PROSITE, PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
 DR PROSITE, PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Multigene family.
 FT DOMAIN 1 26
 FT TRANSMEM 27 47
 FT DOMAIN 48 58
 FT TRANSMEM 59 79
 FT DOMAIN 80 96
 FT TRANSMEM 97 117
 FT DOMAIN 118 138
 FT TRANSMEM 139 159
 FT DOMAIN 160 189
 FT TRANSMEM 190 210
 FT DOMAIN 211 287
 FT TRANSMEM 288 308
 FT DOMAIN 309 321
 FT TRANSMEM 322 342
 FT DOMAIN 343 371
 FT DISULFID 95 173
 FT CARBOHYD 3 3
 FT CARBOHYD 84 84
 FT CARBOHYD 183 183
 SQ SEQUENCE 371 AA; 41954 MW; F6F6175BD3A348C2 CRC64;
 Query Match 68.5%; Score 37; DB 1; Length 371;
 Best Local Similarity 83.3%; Pred. No. 24;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACYKRV 6
 DB 305 ACYKRV 310
 RESULT 8
 SRB3 HUMAN STANDARD; PRT; 373 AA.
 AC Q9NS66;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Super conserved receptor expressed in brain 3.
 GN SRB3.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20294882; PubMed=10833454;
 RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,
 RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuchi K.,
 RT "An evolutionarily conserved G-protein coupled receptor family, SRBB,
 RT expressed in the central nervous system.";
 RL Biochem. Biophys. Res. Commun. 272:576-582 (2000).
 CC
 CC [2]
 CC SEQUENCE FROM N.A.
 CC TISSUE=Lung;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Wozny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL
 CC -1- FUNCTION: ORPHAN RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in brain and ovary.
 CC Lower levels in small intestine. In brain regions, detected in all
 CC regions tested. Highest levels in the cerebellum and cerebral
 CC cortex.
 CC
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL, AB040801; BAA96647.1; -;
 CC EMBL, BC009861; AAH09861.1; -;
 DR PIR, JC7289; JC7289.
 DR MIM, 300253; -;
 DR GO, GO:0007165; P-signal transduction; TNS.
 DR InterPro, IPR000276; GPCR_Rhodopsin.
 DR Pfam, PF00001; 7tm_1, 1.
 DR PRINTS, PR00237; GPCR_Rhodopsin.
 DR PROSITE, PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
 DR PROSITE, PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Multigene family.
 FT DOMAIN 1 26
 FT TRANSMEM 27 47
 FT DOMAIN 48 59
 FT TRANSMEM 60 80
 FT DOMAIN 81 97
 FT TRANSMEM 98 118
 FT DOMAIN 119 139
 FT TRANSMEM 140 160
 FT DOMAIN 161 188
 EXTRACELLULAR (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 2 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 3 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 4 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 189 209 5 (POTENTIAL).
FT DOMAIN 210 287 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 288 308 6 (POTENTIAL).
FT DOMAIN 309 322 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 323 343 7 (POTENTIAL).
FT DOMAIN 344 373 CYTOPLASMIC (POTENTIAL).
FT DISULFID 96 174 BY SIMILARITY.
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 373 AA; 41481 MW; 8A227F914C9D8358 CRC64;

Query Match
Best Local Similarity 68.5%; Score 37; DB 1; Length 373;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTWV 6
DB 305 ACTWV 310

RESULT 9
SRB3_RAT
ID SRB3_RAT STANDARD; PRT; 373 AA.
AC Q9JUH2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Super conserved receptor expressed in brain 3.
GN SREB3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20294882; PubMed=10833454;
RA Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,
RA Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuchi K.;
RT "An evolutionarily conserved G-protein coupled receptor family, SREB,
RT expressed in the central nervous system."
RT Biochem. Biophys. Res. Commun. 272:576-582 (2000).
RL Biochem. Biophys. Res. Commun. 272:576-582 (2000).
CC -1- FUNCTION: ORPHAN RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB040804; BAA96650.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECP F1.1; FALSE_NEG.
DR PROSITE; PS0262; G-PROTEIN RECP F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 27 47 1 (POTENTIAL).
FT DOMAIN 48 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 80 2 (POTENTIAL).
FT DOMAIN 81 97 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 98 118 3 (POTENTIAL).
FT DOMAIN 119 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 4 (POTENTIAL).
FT DOMAIN 161 188 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 189 209 5 (POTENTIAL).
FT DOMAIN 210 287 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 288 308 6 (POTENTIAL).
FT DOMAIN 309 322 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 323 343 7 (POTENTIAL).
FT DOMAIN 344 373 CYTOPLASMIC (POTENTIAL).
FT DISULFID 96 174 BY SIMILARITY.
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 373 AA; 41511 MW; C06DEA2F0E86CAF5 CRC64;

Query Match
Best Local Similarity 68.5%; Score 37; DB 1; Length 373;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTWV 6
DB 305 ACTWV 310

RESULT 10
VNN2_HUMAN
ID VNN2_HUMAN STANDARD; PRT; 520 AA.
AC O95498; Q9UUF3; Q9UWM2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vascular non-inflammatory molecule 2 precursor (Vanin 2)
DE (Glycosylphosphatidyl inositol-anchored protein GPI-80) (FOAP-4
DE protein).
GN VNN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=99009335; PubMed=9790769;
RA Galland F., Malergue F., Bazin H., Mattei M.-G., Aurrand-Lions M.,
RA Theillet C., Naquet P.;
RT "Two human genes related to murine vanin-1 are located on the long arm
RT of human chromosome 6."
RT Genomics 53:203-213 (1998).
RL Genomics 53:203-213 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood;
RX MEDLINE=99218457; PubMed=10201959;
RA Suzuki K., Matanabe T., Sakurai S.-I., Onake K., Kinoshita T.,
RA Araki A., Fujita T., Takei H., Takeda Y., Sato Y., Yamashita T.,
RA Araki Y., Sando F.;
RT "A novel glycosylphosphatidyl inositol-anchored protein on human
RT leukocytes: a possible role for regulation of neutrophil adherence and
RT migration."
RT J. Immunol. 162:4277-4284 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Takayama K., Fujii Y., Tsuritani K., Yajima Y., Amemiya T., Ukai Y.,
RA Naito K., Kawaguchi A.;
RT "Homo sapiens mRNA for FOAP-4 protein, complete cds."
RT Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Phillips S.;
RT Submitted (Feb-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PROBABLE HYDROLASE. INVOLVED IN THE THYMUS HOMING OF
CC BONE MARROW CELLS. MAY REGULATE BETA-2 INTEGRIN-MEDIATED CELL
CC ADHESION, MIGRATION AND MOTILITY OF NEUTROPHIL.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN, THYMUS, PERIPHERAL BLOOD
CC LYMPHOCYTES, AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE CN HYDROLASE FAMILY. BTD/VNN SUBFAMILY.
CC -1- SIMILARITY: Contains 1 CN hydrolase domain.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AJ132100; CA10569.1; -
DR EMBL: AL032821; CAB40076.1; -
DR EMBL: AB026705; BAB61019.1; -
DR EMBL: D89974; BAA82525.1; -
DR Genew: HGNC:12706; VNN2.
DR MIM: 603571; -
DR GO: GO:0006928; P:cell motility; TAS.
DR InterPro: IPR003010; Ntase/CNhydase.
DR Pfam: PF00795; CN_hydrolase; 1.
DR PROSITE: PS50263; CN_HYDROLASE; 1.
KW Hydrolase; Signal; Glycoprotein; GPI-anchor.
FT SIGNAL 1 22
FT CHAIN 1 493
FT PROPEP 494 520
FT DOMAIN 26 328
FT DOMAIN 393 396
FT LIPID 493 493
FT CARBOHYD 39 39
FT CARBOHYD 273 273
FT CARBOHYD 347 347
FT CARBOHYD 357 357
FT CARBOHYD 411 411
FT CARBOHYD 468 468
FT CONFLICT 218 218
FT CONFLICT 404 404
SQ SEQUENCE 520 AA; 58521 MW; B460F1DE57755C6 CRC64;

Query Match 68.5%; Score 37; DB 1; Length 520;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWKVCT 8
DB 398 YWQVCT 403

RESULT 11
MUC2_RAT
ID MUC2_RAT STANDARD; PRT; 1513 AA.
AC 062635;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mucin 2 precursor (Intestinal mucin 2) (Fragment).
GN MUC2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=94299489; Pubmed=8027037;
RA Ohmori H., Dohman A.F., Gallup M., Tauda T., Kai H., Gum J.R. Jr.,
RA Kim Y.S., Babbum C.B.;
RT "Molecular cloning of the amino-terminal region of a rat MUC 2 mucin
RT gene homologue. Evidence for expression in both intestine and
RT airway".
RL J. Biol. Chem. 269:17833-17840(1994).
RN [2]
RP SEQUENCE OF 21-36.
RC TISSUE=Intestinal epithelium;
RX MEDLINE=98180965; Pubmed=9512496;
RA Khatri I.A., Forstner G.G., Forstner J.F.;

RT "Susceptibility of the cysteine-rich N-terminal and C-terminal ends of
RT rat intestinal mucin Muc 2 to proteolytic cleavage".
RL Biochem. J. 331:323-330(1999).
CC -1- FUNCTION: Coats the epithelia of the intestines, airways, and
CC other mucous membrane-containing organs. Thought to provide a
CC protective, lubricating barrier against particles and infectious
CC agents at mucosal surfaces.
CC -1- SUBUNIT: Multimeric.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed in intestine and airway.
CC -1- SIMILARITY: Contains 1 VWF domain.
CC -1- SIMILARITY: Contains 3 VWF domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U07615; AAA21655.2; -
DR PIR: A54895; A54895.
DR InterPro: IPR01007; VWF_C.
DR InterPro: IPR01846; VWF_D.
DR Pfam: PF01826; TIL; 1.
DR Pfam: PR00094; vwd; 3.
DR SMART: SM00214; VWC; 1.
DR SMART: SM00216; VWD; 3.
DR PROSITE: PS01208; VWF_C_1; FALSE NEG.
DR PROSITE: PS0184; VWF_C_2; FALSE NEG.
KW Signal; Repeat; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 1 20
FT DOMAIN 21 1513
FT DOMAIN 21 182
FT DOMAIN 350 410
FT DOMAIN 377 541
FT DOMAIN 847 1005
FT DOMAIN 1392 1513
FT REPEAT 1392 1513
FT REPEAT 1408 1423
FT REPEAT 1424 1434
FT REPEAT 1435 1445
FT REPEAT 1446 1456
FT REPEAT 1457 1467
FT REPEAT 1468 1478
FT REPEAT 1479 1489
FT REPEAT 1490 1500
FT REPEAT 1501 1511
FT REPEAT 1512 1513
FT CARBOHYD 160 160
FT CARBOHYD 420 420
FT CARBOHYD 667 667
FT CARBOHYD 767 767
FT CARBOHYD 837 837
FT CARBOHYD 892 892
FT CARBOHYD 1136 1136
FT CARBOHYD 1151 1151
FT CARBOHYD 1212 1212
FT CARBOHYD 1227 1227
FT CARBOHYD 1243 1243
FT CARBOHYD 1350 1350
FT NON TER 1513 1513
SQ SEQUENCE 1513 AA; 166036 MW; 26109DCA1BE7D008 CRC64;

Query Match 68.5%; Score 37; DB 1; Length 1513;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWVC 7
DB 1235 CYWVC 1240


```

RESULT 12
UR2A CATCO STANDARD; PRT; 12 AA.
AC P04558;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Urotensin IIA (U-IIA) (UIIA).
OS Catestomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catestomidae; Catestomus.
OX NCBI_Taxid=7971;
RN [1]
RP SEQUENCE.
RX MEDLINE=8401959; PubMed=6138758;
RA McMaster D., Lederis K.;
RT "Isolation and amino acid sequence of two urotensin II peptides from
RT Catestomus commersoni urophyses.";
RL Peptides 4:367-373 (1983)
CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY
CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR.
CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
DR PIR: J50423; J50423.
DR InterPro: IPR001483; Urotensin_II.
DR Pfam: PF02083; Urotensin_II; 1.
DR PROSITE: PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID
SQ SEQUENCE 12 AA; 1336 MW; 969C76DBB879CEBA CRC64;

Query Match 66.7%; Score 36; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
Db 6 CFMKYC 11

RESULT 13
UR2B CATCO STANDARD; PRT; 12 AA.
AC P04559;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Urotensin IIB (U-IIB) (UIIB).
OS Catestomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catestomidae; Catestomus.
OX NCBI_Taxid=7971;
RN [1]
RP SEQUENCE.
RX MEDLINE=8401959; PubMed=6138758;
RA McMaster D., Lederis K.;
RT "Isolation and amino acid sequence of two urotensin II peptides from
RT Catestomus commersoni urophyses.";
RL Peptides 4:367-373 (1983)
CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY
CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR.
CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
DR PIR: J50424; J50424.
DR InterPro: IPR001483; Urotensin_II.
DR Pfam: PF02083; Urotensin_II; 1.
DR PROSITE: PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID
SQ SEQUENCE 12 AA; 1437 MW; 73961BDBB879CEBB CRC64;

```

```

Query Match 66.7%; Score 36; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
Db 6 CFMKYC 11

RESULT 14
UR2B CYPCA STANDARD; PRT; 12 AA.
AC P04561;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Urotensin II-Beta.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_Taxid=7962;
RN [1]
RP SEQUENCE.
RA Munekata E., Ohtaki T., Ichikawa T., McMaster D., Lederis K.;
RL (In) Rich D.H., Gross E. (eds.);
RL Proceedings of the 7th American peptide symposium, pp. 69-72,
RL Pierce Chemical Co., Rockford IL. (1981).
CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY
CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR.
CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
DR InterPro: IPR001483; Urotensin_II.
DR Pfam: PF02083; Urotensin_II; 1.
DR PROSITE: PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID
FT VARIANT 6 11 G->S.
FT VARIANT 2 2
SQ SEQUENCE 12 AA; 1407 MW; 73960A9FB879CEBB CRC64;

Query Match 66.7%; Score 36; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 7
Db 6 CFMKYC 11

RESULT 15
UR2 GILMI STANDARD; PRT; 12 AA.
AC P0147;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Urotensin II (U-II) (UII).
OS Gillichthys mirabilis (Long-jawed mudsucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Gobioidae;
OC Gobiidae; Gillichthys.
OX NCBI_Taxid=8222;
RN [1]
RP SEQUENCE.
RX MEDLINE=81054904; PubMed=6107911;
RA Pearson D., Shively J.E., Clark B.R., Gieschwind I.I., Barkley M.,
RA Nishio K., Bern H.A.;
RT "Urotensin II: a somatostatin-like peptide in the caudal
RT neurosecretory system of fishes.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:5021-5024 (1980).
CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY

```

CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR.
CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
DR PIR; A01409; UOGM2.
DR PIR; S42765; S42765.
DR InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KM Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1364 MW; 968BF8982679CEBA CRC64;
Query Match 66.7%; Score 36; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.5;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 CYMKVC 7
|:|:|
DB 6 CFMKYC 11

Search completed: February 11, 2004, 06:32:59
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 06:30:18 ; Search time 34 Seconds
(without alignments)
60.718 Million cell updates/sec

Title: SEQ1
Perfect score: 54
Sequence: 1 acykwvct 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP virus: *
16: SP bacteriap: *
17: SP archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	79.6	468	5	Q81619
2	41	75.9	826	16	Q8E1J1
3	40	74.1	192	12	Q91B09
4	40	74.1	192	12	Q9E228
5	40	74.1	192	12	Q99G02
6	40	74.1	192	12	Q8V505
7	40	74.1	322	16	Q8E1I9
8	40	74.1	335	16	Q8YFA7
9	40	74.1	338	16	Q8G2L2
10	40	74.1	821	5	Q9UAT5
11	40	74.1	821	5	Q966D4
12	39	72.2	500	5	Q95YK5
13	38	70.4	310	16	Q8BX92
14	38	70.4	313	16	Q8P659
15	38	70.4	316	2	Q8GQ67
16	38	70.4	323	2	Q8KWR0

17	38	70.4	381	2	Q9X356	Q9X356 bacillus an
18	38	70.4	452	11	Q8CGA0	Q8CGA0 mus musculu
19	38	70.4	501	5	Q8MSQ0	Q8MSQ0 ciona intes
20	38	70.4	503	5	Q9NAR8	Q9NAR8 branchiosto
21	37	68.5	31	13	Q9DEW1	Q9DEW1 coturnix co
22	37	68.5	370	4	Q8BN22	Q8BN22 homo sapien
23	37	68.5	435	3	Q8NML7	Q8NML7 aspergillus
24	37	68.5	773	13	Q8JHW0	Q8JHW0 rana ridibu
25	37	68.5	1567	10	Q80645	Q80645 arabidopsis
26	37	68.5	2994	5	Q811Z4	Q811Z4 plasmodium
27	36	66.7	85	6	Q95K72	Q95K72 sus scrofa
28	36	66.7	121	6	Q95J46	Q95J46 sus scrofa
29	36	66.7	125	6	Q8HYC2	Q8HYC2 macaca mula
30	36	66.7	253	10	Q9M1N1	Q9M1N1 arabidopsis
31	36	66.7	394	5	Q8SVG5	Q8SVG5 encephalito
32	36	66.7	490	11	Q8K0G0	Q8K0G0 mus musculu
33	36	66.7	498	11	Q9D001	Q9D001 mus musculu
34	36	66.7	555	4	Q96K49	Q96K49 homo sapien
35	36	66.7	555	11	Q8BKU8	Q8BKU8 mus musculu
36	36	66.7	627	10	Q9ZV83	Q9ZV83 arabidopsis
37	36	66.7	745	5	Q95X11	Q95X11 caenorhabdi
38	36	66.7	765	11	Q8C1E4	Q8C1E4 mus musculu
39	36	66.7	1375	3	Q9UTL9	Q9UTL9 schizosacch
40	36	66.7	1487	5	Q9VLP5	Q9VLP5 drosophila
41	35.5	65.7	623	12	Q8V0N4	Q8V0N4 planococcus
42	35	64.8	69	16	Q9CG77	Q9CG77 lactococcus
43	35	64.8	171	12	Q9PYR0	Q9PYR0 xestia c-ni
44	35	64.8	230	17	Q8PXG1	Q8PXG1 methanosarc
45	35	64.8	238	16	Q8EX27	Q8EX27 mycoplasma

ALIGNMENTS

RESULT 1	Q81619	PRELIMINARY;	PRT;	468 AA.
ID	Q81619			
AC	Q81619;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Zic family transcription factor.			
GN	HRZ1CN.			
OS	Halocynthia roretzi (Sea squirt).			
OC	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;			
OC	Stolidobranchia; Pyuridae; Halocynthia.			
OX	NCBI_Taxid=7729;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22308874; PubMed=12421701;			
RA	Wada S., Saiga H.;			
RT	"Hrzn1", a new Zic family gene of ascidians, plays essential roles in			
RT	the neural tube and notochord development.";			
RT	Development 129:5597-5608 (2002).			
DR	EMBL; AB092643; BAC23063.1; -			
SO	SEQUENCE	468 AA;	51632 MW;	DBS67BA5C1PF31 CRC64;
Query Match		79.6%;	Score 43;	DB 5;
Best Local Similarity		85.7%;	Pred. No. 10;	Length 468;
Matches	6;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;
QY	2	CYWKVCT 8		
DB	158	CYWKDCT 164		
RESULT 2	Q8E1J1	PRELIMINARY;	PRT;	826 AA.
ID	Q8E1J1			
AC	Q8E1J1;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			

DE Periplasmic nitrate reductase.
 GN NAPA OR S00848.
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 OX NCBI_Taxid=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MR-1;
 RX MEDLINE=22297686; PubMed=12368813;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Meche B., Clayton R.A.,
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
 RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
 RA Fouts D.E., Kouri H., Gill J., Utterback T.R., McDonald L.A.,
 RA Fiedlitz T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.,
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
 RT Shewanella oneidensis.";
 RL Nat. Biotechnol. 20:1118-1123 (2002).
 DR EMBL; AF015529; AAN53924.1; -.
 DR TIGR; S00848; -.
 KW Complete proteome.
 SQ SEQUENCE 826 AA; 92474 MW; 33474D23868F6316 CRC64;

Query Match 75.9%; Score 41; DB 16; Length 826;
 Best Local Similarity 71.4%; Pred. No. 37;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVC 8
 DB 474 CWTMCT 480

RESULT 3
 Q91B09 PRELIMINARY; PRT; 192 AA.
 ID Q91B09;
 AC Q91B09;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Helicoverpa armigera nuclear polyhedrosis virus.
 OS Helicoverpa armigera nuclear polyhedrosis virus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_Taxid=51313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CI;
 RX PubMed=12050807;
 RA Zhang C.X., Wu J.C.;
 RT "Genome structure and the p10 gene of the Helicoverpa armigera
 RT nucleopolyhedrovirus.";
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 33:179-184 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CI;
 RX Zhang C.X., Jin W.R.;
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF303045; AAK96351.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 192 AA; 22745 MW; 94138789371F641E CRC64;

Query Match 74.1%; Score 40; DB 12; Length 192;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 DB 61 CHMKVC 66

RESULT 4
 Q9E228 PRELIMINARY; PRT; 192 AA.
 ID Q9E228;
 AC Q9E228;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 22.7 kDa protein.
 OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_Taxid=10468;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Robertson A.P.S.;
 RT "Genetic organization of Helicoverpa zea nuclear polyhedrosis virus in
 RT the region of Scott U.D.L.A and Q.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF275264; MAG17374.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 192 AA; 22747 MW; D5138789322580B5 CRC64;

Query Match 74.1%; Score 40; DB 12; Length 192;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 DB 61 CHMKVC 66

RESULT 5
 Q99G02 PRELIMINARY; PRT; 192 AA.
 ID Q99G02;
 AC Q99G02;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 22.7 kDa protein.
 OS Helicoverpa armigera nucleopolyhedrovirus G4.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_Taxid=148363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Deng F., Chen X., Vlak J.M., Arif B.M., Hu Z.;
 RT "Sequence analysis of the gp37 gene of Heliothis armigera single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RL Zhongguo Bingduxue 15:35-42 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wang H., Hu Z., Sun X., Vlak J.M., Chen X.;
 RT "Sequence analysis of the Iap3 gene of Heliothis armigera single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RL Zhongguo Bingduxue 15:43-49 (2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21078302; PubMed=11210934;
 RA Wang H., Chen X., Wang H., Arif B.M., Vlak J.M., Hu Z.;
 RT "Nucleotide sequence and transcriptional analysis of a putative basic
 RT DNA-binding protein of Helicoverpa armigera polyhedrovirus.";
 RL Virus Genes 22:113-120 (2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21064569; PubMed=11125177;
 RA Chen X., Ikel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
 RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
 RT "The sequence of the Helicoverpa armigera single-nucleocapsid
 RT nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 82:241-257 (2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Chen X., Ikel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,

RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF271059; AAC53851.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 192 AA; 22773 MW; 941387892FBEED1E CRC64;

Query Match 74.1%; Score 40; DB 12; Length 192;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 DB 61 CHMKVC 66

RESULT 6
 Q8V505 PRELIMINARY; PRT; 192 AA.
 AC Q8V505;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE ORF11.
 OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OK NCBI_TaxID=10468;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chen X., Zhang W.-J., Wong J., Chun G., Lu A., McCutchen B.F.,
 RA Prenatal J.K., Herrmann R., Dolan M., Tingey S., Hu Z.-H., Vlak J.M.;
 RT "Genome sequence analysis of Helicoverpa zea single nucleocapsid
 RT nucleopolyhedrovirus." to the EMBL/GenBank/DBJ databases.
 DR Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A334030; AAL56117.1; -;
 SQ SEQUENCE 192 AA; 22675 MW; 549DBCE484FEE8D CRC64;

Query Match 74.1%; Score 40; DB 12; Length 192;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYMKVC 7
 DB 61 CHMKVC 66

RESULT 7
 Q8Z119 PRELIMINARY; PRT; 322 AA.
 AC Q8Z119;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein STY4576.
 GN STY4576.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OK NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CT18;
 RA MEDLINE-21534947; PubMed-11677608;
 RA Parhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Hogue A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Mould S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Skellern M., Skellern J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";

RL Nature 413:848-852(2001).
 DR EMBL; AL627282; CAD09351.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 322 AA; 34841 MW; 8E08C95F19237648 CRC64;

Query Match 74.1%; Score 40; DB 16; Length 322;
 Best Local Similarity 71.4%; Pred. No. 24;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVCT 8
 DB 49 CYMLMCT 55

RESULT 8
 Q8YFA7 PRELIMINARY; PRT; 335 AA.
 AC Q8YFA7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transcriptional regulator, ARAC family.
 GN BME1615.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OK NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kaputal V., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,
 RA Jabinok E., Elizer P.H., Hagius S., D'Souza M., Bernal A., Mazur M.,
 RA Selkov E., Elizer P.H., Hagius S., D'Souza M., Bernal A., Mazur M.,
 RA Haselkorn R., Kyprides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis." to the EMBL/GenBank/DBJ databases.
 DR Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL; AE009597; AAL52796.1; -;
 DR InterPro; IPR000005; HTHARAC.
 DR InterPro; IPR002818; TH1J.
 DR Pfam; PR01965; D2-1_Pf01.1.
 DR Pfam; PR00165; HTH_ARAC.2.
 DR SMART; SM00342; HTH_ARAC.1.
 DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 335 AA; 37076 MW; F66631F5217F38D6 CRC64;

Query Match 74.1%; Score 40; DB 16; Length 335;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYMKVCT 8
 DB 60 CAMKICT 66

RESULT 9
 Q8G2L2 PRELIMINARY; PRT; 338 AA.
 AC Q8G2L2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transcriptional regulator, Arac family.
 GN BR0307.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OK NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=2247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AE014343; AAN29256.1; -.
 DR TIGR; BR0307; -.
 KW Complete proteome.
 SO SEQUENCE 338 AA; 37205 MW; 915F364BACE15C5 CRC64;

Query Match 74.1%; Score 40; DB 16; Length 338;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 YMKVCT 8
 DB 63 YMKICT 69

RESULT 10

O9UAT5 PRELIMINARY; PRT; 475 AA.

AC O9UAT5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 53.1 kDa protein.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodexinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RX MEDLINE=9069613; PubMed=9851916;

RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RA Smith A., Wameley P., Fronick W.;
 RT "The sequence of C. elegans cosmid C01B4.";
 RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A0128552; AAD14699.1; -.
 DR WormPep; C01B4.7; CE20476.
 DR InterPro; IPR007114; MFS.
 DR PROSITE; PS50850; MFS. 1.
 KW Hypothetical protein.
 SO SEQUENCE 475 AA; 53094 MW; 79095D45572AF535 CRC64;

Query Match 74.1%; Score 40; DB 5; Length 475;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 YMKVCT 8
 DB 264 YMKICT 269

RESULT 11

O966D4 PRELIMINARY; PRT; 821 AA.

AC O966D4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein Y19D10A.4.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodexinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RX MEDLINE=9069613; PubMed=9851916;

RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RA Johnson D.;
 RT "The sequence of C. elegans cosmid Y19D10A.";
 RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006722; AAK68417.1; -.
 DR WormPep; Y19D10A.4; CE21450.
 DR InterPro; IPR001823; Ald1_epimerase.
 DR InterPro; IPR007114; MFS.
 DR Pfam; PF01263; Aldose epim; 1.
 DR PROSITE; PS50850; MFS. 1.

SO SEQUENCE 821 AA; 91594 MW; 923A788FC95D1A76 CRC64;

Query Match 74.1%; Score 40; DB 5; Length 821;
 Best Local Similarity 83.3%; Pred. No. 56;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 YMKVCT 8
 DB 264 YMKICT 269

RESULT 12

O95YK5 PRELIMINARY; PRT; 500 AA.

AC O95YK5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Zic related protein 2.

OS Ciona savignyi.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Clonidae; Ciona.
 OX NCBI_TaxID=51511;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Satou Y., Imai K.S., Satoh N.;
 RT "Ciona savignyi genes.";
 RN Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB057740; BAB68349.1; -.
 DR InterPro; IPR007087; ZnF_C2H2.
 DR Pfam; PF00096; zf_C2H2; 4.
 DR Prodom; PD000003; ZnF_C2H2; 1.

DR SMART: SM00355; ZNF C2H2; 4.
 DR PROSITE: PS00026; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 4.
 KW Metal-binding; zinc; Zinc-finger.
 SQ SEQUENCE 500 AA; 55771 MW; 9A38B537017BFP3F1 CRC64;

Query Match 72.4%; Score 39; DB 5; Length 500;
 Best Local Similarity 62.5%; Pred. No. 53;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACYWKVCT 8
 Db 183 ACYWKVCT 190

RESULT 13
 ID 08PK92 PRELIMINARY; PRT; 310 AA.
 AC 08PK92;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein XAC2283.
 GN XAC2283.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OC NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).
 DR EMBL: AB011865; AAM37136.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 310 AA; 33451 MW; A9781D41C3AC9C55 CRC64;

Query Match 70.4%; Score 38; DB 16; Length 310;
 Best Local Similarity 71.4%; Pred. No. 51;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
 Db 48 CYWKVCT 54

RESULT 14
 ID 08P659 PRELIMINARY; PRT; 313 AA.
 AC 08P659;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein XCC3115.
 GN XCC3115.
 OS Xanthomonas campestris (pv. campestris).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OC NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavan F., Cardoso J., Chamberg F., Clapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).
 DR EMBL: AB012427; AAM42386.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 313 AA; 33935 MW; 3599D91640226CE8 CRC64;

Query Match 70.4%; Score 38; DB 16; Length 313;
 Best Local Similarity 71.4%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
 Db 50 CYWKVCT 56

RESULT 15
 ID 08G067 PRELIMINARY; PRT; 316 AA.
 AC 08G067;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN Pseudomonas aeruginosa.
 OS Pseudomonadaceae; Pseudomonas.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C;
 RX MEDLINE=22313472; PubMed=12426355;
 RA Larbig K.D., Christmann A., Johann A., Klockgether J., Hartsch T.,
 RA Merkl R., Wihlmann L., Fritz H.J., Jummeler B.,
 RT "Gene Islands Integrated into tRNA(Gly) Genes Confer Genome Diversity
 RT on a Pseudomonas aeruginosa Clone."
 RL J. Bacteriol. 184:6665-6680(2002).
 DR EMBL: AF440523; AAM62136.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 316 AA; 34725 MW; D2FF28A4C843C3E0 CRC64;

Query Match 70.4%; Score 38; DB 2; Length 316;
 Best Local Similarity 71.4%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
 Db 53 CYWKVCT 59

Search completed: February 11, 2004, 06:32:38

Job time : 38 secs

DR WPI; 1990-172518/23.
 XX Somatostatin and its analogues or derivs. - used for protection against
 PT lung damage, esp. for respiratory distress syndrome or damage caused
 PT by oxygen therapy.
 XX
 PS Disclosure; Page 7; 7pp; German.
 XX
 CC Residues A at positions 1 is beta-Naphthyl-(D)Alanine and residue
 CC Trp at position 4 is D-Trp.
 CC Derivs of this sequence substd. on the N-terminal amino gp. by
 CC N alpha-(alpha-glucosyl(1-4)deoxyfructosyl)
 CC or N alpha-(beta-deoxyfructosyl) are pref. for oral administration.
 CC The cpds. are esp. used to prevent, inhibit or treat respiratory
 CC distress syndrome in children and adults, and to inhibit or treat lung
 CC damage caused by oxygen therapy. Opt. they are formulated with a
 CC glucocorticoid steroid. The y are also known to inhibit release of
 CC growth hormone, insulin, and glucagon, and to reduce gastric secretions.
 CC See also AAR05177-80, AAR05182-88, AAR05191-95, AAR06530 and AAR05197.
 CC (Updated on 10-MAR-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC
 SQ Sequence 8 AA;
 XX
 Query Match 100.0%; Score 54; DB 11; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACYWKVCT 8
 DB 1 ACYWKVCT 8
 XX
 RESULT 2
 AAR41505
 ID AAR41505 standard; peptide; 8 AA.
 XX
 AC AAR41505;
 XX
 DT 25-MAR-2003 (updated)
 DT 23-FEB-1994 (first entry)
 XX
 DE Somatostatin peptide analogue f.
 XX
 KM Somatostatin; peptide analogue; growth hormone; secretion; pancreas;
 KM gastric secretion; acromegaly; diabetes mellitus; pituitary tumour;
 KM gastro-intestinal disorder; headache; migraine.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /label= OTHER
 FT /note= "3-(2-(naphthyl)-(D)Ala"
 FT Disulfide-bond 2..7
 FT Misc-difference 4 /note= "D-form residue"
 FT Modified-site 8 /label= Thr-NH2
 FT
 XX
 PN W09317037-A1.
 XX
 PD 02-SEP-1993.
 XX
 PF 16-FEB-1993; 93WO-EP00366.
 XX
 PR 21-FEB-1992; 92GB-0003769.
 XX
 PA (SANO) SANDOZ ERFINDEUNGEN VERW GMBH.
 PA (SANO) SANDOZ LTD.
 PA (SANO) SANDOZ PATENT GMBH.
 XX
 PI Hirt D, Latasec X;

XX
 DR WPI; 1993-208354/36.
 XX
 PT Treating acute migraine or cluster headache - with somatostatin
 PT peptide, pref. admin. as nasal powder, providing rapid relief
 XX
 PS Claim 5; Page 24; 29pp; English.
 XX
 CC The sequences given in AAR41500-08 are somatostatin peptide analogues.
 CC These peptides are based on sequences from the tetradecapeptide
 CC somatostatin but have deletions, substitutions or inclusions when
 CC compared to the wild type sequence. These sequences may be used in
 CC a compound to inhibit growth hormone secretion and pancreatic and
 CC gastric secretion. This compound may be used in the treatment of
 CC acromegaly, complications of diabetes mellitus or gastro-intestinal
 CC disorders. These peptides may also be used to alleviate chronic,
 CC life-disrupting headache associated with pituitary tumours and also
 CC migraine.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 SQ Sequence 8 AA;
 XX
 Query Match 100.0%; Score 54; DB 14; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACYWKVCT 8
 DB 1 ACYWKVCT 8
 XX
 RESULT 3
 AAW97188
 ID AAW97188 standard; peptide; 8 AA.
 XX
 AC AAW97188;
 XX
 DT 05-MAY-1999 (first entry)
 XX
 DE Somatostatin analog peptide sequence.
 XX
 KM Somatostatin; sustained-release; microparticle; octreotide;
 KM somatostatin analog; calcitonin; tyressin; growth hormone secretion;
 KM acromegaly; breast cancer; gastrointestinal disorder; peptic ulcer;
 KM enterocutaneous; pancreaticocutaneous fistula; irritable bowel syndrome;
 KM dumping syndrome; watery diarrhea syndrome; acute pancreatitis;
 KM gastroenteropathic endocrine tumor; glucagonoma; insulinoma; gastrinoma;
 KM carcinoid tumour; gastrointestinal bleeding; anticancer; neuroregulator;
 KM growth hormone inhibitor.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "3-(2-Naphthyl) attached to D-form residue"
 FT Misc-difference 4 /note= "D-form residue"
 FT Modified-site 8 /note= "amidated residue"
 FT
 XX
 PN US5876761-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 06-JUN-1995; 95US-0470907.
 XX
 PR 18-JAN-1991; 91US-0643880.
 PR 07-JUL-1989; 89US-0377023.
 PR 22-SEP-1989; 89US-0411347.
 PR 06-JUN-1995; 95US-0470907.
 XX
 PA (NOVS) NOVARTIS AG.
 PA
 XX

PI Bodmer D, Fong JW, Kissel T, Maulding HV, Nagele O;
 PI Pearson JE;
 DR MPI; 1999-189594/16.
 XX
 XX
 PT Production of microparticles comprising octreotide - by mixing
 PT octreotide solution in buffer and polylactide:glycolide solution
 PT in methylene chloride, used to treat breast cancer or acromegaly
 XX
 PS Disclosure; Column 5; 9pp; English.
 XX
 CC The present sequence represents a somatostatin analog that can be used
 CC in the production of sustained-release microparticles comprising
 CC octreotide. The microparticles are used for sustained-release
 CC formulations of drugs, particularly water soluble peptides e.g.
 CC somatostatin or somatostatin analogs, such as octreotide, as well as
 CC calcitonin (e.g. salmon calcitonin) and tyressin. Microparticles
 CC produced are used in the long-term treatment of disorders associated
 CC with excess growth hormone secretion such as acromegaly, breast cancer
 CC and gastrointestinal disorders such as peptic ulcer, enterocutaneous and
 CC pancreaticocutaneous fistula, irritable bowel syndrome, dumping syndrome,
 CC watery diarrhea syndrome, acute pancreatitis, gastroenteropathic
 CC endocrine tumors (e.g. VIPomas, GRPomas, glucagonomas, insulinomas,
 CC gastrinomas and carcinoid tumors) and gastrointestinal bleeding, and
 CC diabetic complications. The microparticles are used in anticancer
 CC treatment, neuroregulators, growth hormone inhibitors, central nervous
 CC system inhibitors, peripheral nervous system inhibitors, gastrointestinal
 CC inhibitors, vascular smooth muscle inhibitors, insulin inhibitors and
 CC glucagon inhibitors.
 XX
 SQ Sequence 8 AA;
 XX
 Query Match 100.0%; Score 54; DB 20; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACYMKVCT 8
 |||||
 DB 1 ACYMKVCT 8
 |||||
 RESULT 4
 AAB19027
 ID AAB19027 standard; peptide; 8 AA.
 XX
 AC AAB19027;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Amino acid sequence of somatostatin analogue lanreotide.
 XX
 KW Somatostatin analogue; somatostatin receptor; eye; growth factor;
 KW choroidal neovascularisation related disorder; ocular neovascularisation;
 KW CNV related disease; exudative age-related macular degeneration;
 KW diabetic retinopathy.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "3-(2-(naphthyl) attached to D form residue"
 FT Disulfide-bond 2..7
 FT Misc-difference 4
 FT /note= "D form residue"
 FT Misc-difference 8
 FT /note= "amidated residue"
 FT
 PN EP1040837-A2.
 XX
 PD 04-OCT-2000.
 XX
 PF 28-FEB-2000; 2000EP-0200707.
 XX

PR 26-FEB-1999; 99US-0258240.
 XX
 XX
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Kuipers RWM, Van Hagen PM, Baarsma GS;
 XX
 DR MPI; 2000-629601/61.
 XX
 XX
 PT Use of somatostatin, somatostatin analogs or somatostatin receptor
 PT binding compounds for treating choroidal neovascularization-related
 PT disorder, especially age-related macular degeneration and diabetic
 PT retinopathy -
 XX
 PS Disclosure; Page 5; 16pp; English.
 XX
 CC AAB19022-32 represent somatostatin analogues. They are used in the
 CC method of the invention. The specification describes the use of a
 CC compound that binds to at least one somatostatin receptor in the eye, in
 CC the manufacture of a medicine for the treatment of a choroidal
 CC neovascularisation (CNV) related disorder. The control of ocular
 CC neovascularisation is regulated by a balance between stimulating and
 CC inhibiting growth factors. The medicine is useful for the treatment of
 CC CNV related diseases, especially exudative age-related macular
 CC degeneration. The somatostatin receptor binding compounds stop leakage
 CC in existing and new ocular vessels and inhibit neovascularisation in
 CC age-related macular degeneration and diabetic retinopathy.
 XX
 SQ Sequence 8 AA;
 XX
 Query Match 100.0%; Score 54; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACYMKVCT 8
 |||||
 DB 1 ACYMKVCT 8
 |||||
 RESULT 5
 AAE19365
 ID AAE19365 standard; peptide; 8 AA.
 XX
 AC AAE19365;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Somatostatin analogue peptide #6.
 XX
 KW Somatostatin analogue; ocular disorder; diabetic macular oedema;
 KW retinitis pigmentosa; optic nerve disease; vision loss; glaucoma;
 KW age related macular degeneration.
 XX
 OS Undefined.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "3-(2-Naphthyl)-Ala; D-form residue"
 FT Disulfide-bond 2..7
 FT Misc-difference 4
 FT /note= "D-form residue"
 FT Modified-site 8
 FT /note= "C-terminal amide"
 FT
 PN WO200209739-A1.
 XX
 PD 07-FEB-2002.
 XX
 PF 25-JUL-2001; 2001WO-BP08620.
 XX
 PR 27-JUL-2000; 2000EP-0116068.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX Fsadni M, Myse K;
XX WPI; 2002-227075/28.
XX
XX Treatment of ocular disorder e.g. glaucoma using a somatostatin analog
XX
XX
XX Claim 4f; Page 10; 20pp; English.
XX
XX The invention relates to a method of treating ocular disorder. The
XX method involves administering a somatostatin analogue to an individual.
XX The method is used to treat ocular disorders particularly vision loss
XX due to progressive retinal and/or optic nerve disease such as glaucoma,
XX retinitis pigmentosa, diabetic macular oedema and age related macular
XX degeneration. The present sequence is somatostatin analogue peptide
XX also known as lanreotide.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 54; DB 23; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACYWKVCT 8
Db 1 ACYWKVCT 8

RESULT 6
AAW48638
ID AAW48638 standard; peptide; 7 AA.
XX
XX AAW48638;
XX
XX 04-AUG-1998 (first entry)
XX
XX Somatostatin peptide analog 1.
XX
XX Somatostatin; acromegaly; radiolocalisation; blood-brain barrier;
XX cancer; halogen.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 3 /note= "D-form residue"
XX Modified-site 7 /note= "C-terminal amide"
XX
XX W0969161-A1.
XX
XX 12-DEC-1996.
XX
XX 03-JUN-1996; 96WO-US08437.
XX
XX 05-JUN-1995; 95US-0462223.
XX
XX (CHIL-) CHILDRENS HOSPITAL INC.
XX (TULA) TULANE EDUCATIONAL FUND.
XX (LOUJ) LOUISIANA STATE MEDICAL CENT FOUND.
XX (OHIS) UNIV OHIO STATE RES FOUND.
XX
XX Coy DH, Murphy MA, O'Dorisio MS, O'Dorisio TM, Woltering EA;
XX WPI; 1997-042842/04.
XX
XX Multi-tyrosinated somatostatin analogues - useful for diagnosis and
XX treatment of diseases related to altered somatostatin receptor
XX expression.
XX
XX Claim 2; Page 59; 63pp; English.

XX The invention provides for somatostatin analogues such as the present
XX one. These somatostatin analogues are claimed to be useful for treating
XX disease associated with increased production of factors which can be
XX regulated by somatostatin, e.g. acromegaly. Also when radioactively
XX labelled, the analogues are claimed to be useful for diagnosing cancer
XX in vitro or in situ where aberrant expression of somatostatin receptors
XX is involved. When halogenated, these analogues are found to bind
XX somatostatin receptors with such high affinity that binding is nearly
XX irreversible under physiological conditions. Therefore, by using
XX radioactive halogenated analogues, the invention claims for improved
XX sensitivity of radiolocalisation of the receptors. The
XX multi-tyrosinated somatostatin analogues (AAW48641-W48646) bind to
XX somatostatin receptors with practically the same affinity as native
XX somatostatin. Multi-tyrosinated analogues have increased half-lives in
XX vivo relative to conventional somatostatin analogues, are resistant to
XX enzymatic degradation and have increased blood-brain barrier
XX penetration.
XX
SQ Sequence 7 AA;

Query Match 92.6%; Score 50; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CYWKVCT 8
Db 1 CYWKVCT 7

RESULT 7
AAB91464
ID AAB91464 standard; Peptide; 7 AA.
XX
XX AAB91464;
XX
XX 22-JUN-2001 (first entry)
XX
XX Angiotensin peptide SEQ ID NO:640.
XX
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
XX blood component; modification; succinimidyl; maleimido group; amino;
XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX W0200069900-A2.
XX
XX 23-NOV-2000.
XX
XX 17-MAY-2000; 2000WO-US13576.
XX
XX 17-MAY-1999; 99US-0134406.
XX 10-SEP-1999; 99US-0153406.
XX 15-OCT-1999; 99US-0159783.
XX
XX (CONJ-) CONJUCHEM INC.
XX
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;
XX WPI; 2001-112059/12.
XX
XX Modifying and attaching therapeutic peptides to albumin prevents
XX peptidase degradation, useful for increasing length of in vivo activity
XX
XX
XX Disclosure; Page 407; 733pp; English.
XX
XX The present invention describes a modified therapeutic peptide (I)
XX comprising a therapeutically active amino acid region (III) and a
XX reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
XX a less therapeutically active amino acid region (IV), which covalently

CC bonds with amino/hydroxy/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (1) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity
 CC in vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases.
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention.

XX SQ Sequence 7 AA;

Query Match 92.6%; Score 50; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CYMKVCT 8
 Db 1 CYMKVCT 7

RESULT 8

AAE17083
 ID AAE17083 standard; peptide; 7 AA.

AC AAE17083;

DT 18-APR-2002 (first entry)

XX Angiotensin peptide used in functional assay for human GPCR11 sequence.

KW Human; G-protein coupled receptor; GPCR11; cerebroprotective; vomiting;
 KW receptor-mediated disorder; therapy; urinary retention; allergy; obesity;
 KW osteoporosis; angina pectoris; restenosis; atherosclerosis; hypotension;
 KW anorexia; tumour; migraine; acute heart failure; ulcer; anti-inflammatory;
 KW stroke; hypertension; neuronal disorder; myocardial infarction psychotic;
 KW depression; mental retardation; neurodegenerative disease; antibacterial;
 KW Alzheimer's disease; dementia; ischaemia; Parkinson's disease; antiviral;
 KW Huntington's disease; anxiety; antifungal; immunosuppressive; cytostatic;
 KW vulnery; analgesic; anorectic; anabolic; diuretic; cardiac; nootropic;
 KW antiemetic; vasotrophic; diabetes; cancer; tranquilizer; neuroleptic;
 KW angiotensin.

XX Synthetic.

OS Key Location/Qualifiers

FH Disulfide-bond 1..6

FT Modified-site 1 /note= "Cys is modified with D-Nal"

FT Misc-difference 3 /note= "D-form residue"

FT Modified-site 7 /note= "C-terminal amide"

FT WO200196330-A2.

PN 27-DEC-2001.

XX 20-JUN-2001; 2001WO-BE00104.

XX 20-JUN-2000; 2000US-212913P.

XX 11-JUL-2000; 2000US-217494P.

XX 26-JAN-2001; 2001EP-0870015.

XX 12-FEB-2001; 2001EP-0870024.

XX (EURO-) EUROSCHREIN SA.

XX Lannoy V, Brezillon S, Dethaux M, Parmentier M, Govarts C;

XX

DR WPI; 2002-130789/17.

XX New G-protein coupled receptor, useful in the manufacture of
 PT medicaments for treating receptor mediated disorders e.g. acute heart
 PT failure and Alzheimer's disease -

XX Example 1; Page 15; 46pp; English.

CC The present invention relates to a G-protein coupled receptor (GPCR) and
 CC nucleotide encoding it. GPCR are useful in the manufacture of a
 CC medicament for the prevention and/or treatment of receptor-mediated
 CC disorders e.g. viral infections, virus and bacterial diseases, diseases
 CC and disorders involving disturbances of cell migration, diseases or
 CC perturbations of immune system including cancers, development of tumours
 CC and tumour metastasis, inflammatory and neoplastic processes; bacterial
 CC and fungal infections, in wound and bone healing, dysfunction of
 CC regulatory growth functions; pains, diabetes, obesity, anorexia, bulimia,
 CC urinary retention, osteoporosis, angina pectoris, atherosclerosis,
 CC restenosis, diseases involving excessive or reduced proliferation or loss
 CC of smooth muscle cells, aneurysm, stroke, ischaemia, ulcers, allergies,
 CC benign prostatic hypertrophy, migraine, vomiting, blood circulating
 CC affections including acute heart failure, hypotension, hypertension and
 CC myocardial infarction psychotic; neuronal disorders such as anxiety,
 CC schizophrenia, manic depression, depression, delirium, dementia, severe
 CC mental retardation; degenerative diseases; neurodegenerative diseases
 CC such as Alzheimer's disease, Parkinson's disease; and dyskinesias e.g.
 CC Huntington's disease or Gilles de la Tourette's syndrome and other
 CC related diseases. The present sequence is angiotensin peptide used in
 CC functional assay for human GPCR11 sequence.

XX SQ Sequence 7 AA;

Query Match 92.6%; Score 50; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CYMKVCT 8
 Db 1 CYMKVCT 7

RESULT 9

AAE1464
 ID AAE1464 standard; protein; 8 AA.

AC AAE1464;

DT 25-MAR-2003 (updated)

DT 27-OCT-1991 (first entry)

XX Cyclic octapeptide somatostatin analogue.

KW Growth hormone; prolactin; insulin; glucagon; gastrin; secretin;
 KW Cholecystokinin.

XX Synthetic.

OS Key Location/Qualifiers

FH Modified-site 1 /label= D-Phe

FT Disulfide-bond 2..7

FT Modified-site 4 /label= D-Trp

FT EP203031-A.

XX 26-NOV-1986.

XX 15-APR-1986; 86EP-0810174.

XX 28-MAR-1986; 86US-0843539.

XX 25-APR-1985; 85US-0727105.

XX

Query Match	92.6%	Score 50;	DB 7;	Length 8;
Best Local Similarity	100.0%	Pred. No. 9.3e+05;		
Matches	7;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

RESULT 10
AAP71378
ID AAP71378 standard; peptide; 8 AA.

DT	25-MAR-2003	(updated)
DT	03-OCT-2002	(updated)
DT	03-MAY-1991	(first entry)
XX		
DE	Sequence of somatostatin analogue.	
XX		
KW	Growth hormone secretion inhibitor; cancer therapy;	
KW	Alzheimer's disease; diabetes.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	1
FT	/label= D-beta-Nal	
FT	Misc-difference	4
FT	/label= D-Trp	
FT	Misc-difference	8
FT	/label= Thr-NH2	
XX		
PN	EP214872-A.	
XX		
PD	18-MAR-1987.	
XX		
XX		
PF	12-SEP-1986;	86EP-0307044.
XX		
PR	17-JUN-1986;	86US-0875266.
PR	12-SEP-1985;	85US-0775488.
PR	22-JUN-1988;	88US-0208883.
PR	17-FEB-1988;	89US-0312138.
XX		
PA	(TULSA) TULANE EDUCATIONAL FUND FUND.	
XX		
PI	Coy DH, Murphy WA, Heman ML;	
XX		
DR	WPI; 1987-074173/11.	

Query Match	92.6%;	Score 50;	DB 8;	Length 8;
Best Local Similarity	100.0%;	Pred. No. 9.3e+05;		
Matches	7;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

RESULT 11
AAP71379
ID AAP71379 standard; peptide; 8 AA.

DT	25-MAR-2003	(updated)
DT	03-OCT-2002	(updated)
DT	03-MAY-1991	(first entry)
XX		
DE	Sequence of somatostatin analogue.	
XX		
KW	Growth hormone secretion inhibitor; cancer therapy;	
KM	Alzheimer's disease; diabetes.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	1 /label= pentafluoro-D-Phe
FT	Misc-difference	4 /label= D-Trp
FT	Misc-difference	8 /label= Thr-NH2
XX		
PN	EP214872-A.	
XX		
PD	18-MAR-1987.	
XX		
PF	12-SEP-1986;	86EP-0307044.
XX		
PR	17-JUN-1986;	86US-0875266.
PR	12-SEP-1985;	85US-0775488.
PR	22-JUN-1988;	88US-0209883.
PR	17-FEB-1989;	89US-0312138.
XX		
PA	(TULIA) TULANE EDUCATIONAL FUND FUND.	
XX		
PI	Coy DH, Murphy WA, Heman ML;	
XX		
DR	WPI; 1987-074173/11.	
XX		
PT	Somatostatin analogue octa:peptide(s) with enhanced activity -	
PT	are useful in the retn. of growth hormone, insulin glucagon, or	
PT	pancreatic secretion, and have anticancer activity	
XX		

PS Disclosure; p9; 10pp; English.
 CC Cpd. conf. the claimed peptides can be used instead of
 CC somatostatin. They are administered in somatostatin dosages or
 CC because of their greater potency, in smaller dosages. Dosage is
 CC 0.01-100 mcg./kg./day pref. 0.1-100 mcg./kg./day to mammal eg
 CC human.
 CC (Updated on 03-OCT-2002 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 8 AA;

Query Match 92.6%; Score 50; DB 8; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CYMKVCT 8
 |||||
 DB 2 CYMKVCT 8

RESULT 12

AA07022
 ID AAR07022 standard; protein; 8 AA.

AC AAR07022;

DT 21-JAN-1991 (first entry)

DE Octapeptide somatostatin analogue.

KW Growth hormone dependent cancer; diabetes; Alzheimers disease;
 KM Cirrhosis; hepatitis; insulin; glucagon.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /label= Dextrorotatory
 FT /note= "Pentafluoro-D-Phe"

FT Modified-site 3 /label= Iodinated phenyl ring

FT Modified-site 4 /label= Dextrorotatory

XX EP389180-A.

XX 26-SEP-1990.

XX 15-MAR-1990; 90EP-0302760.

XX 15-MAR-1989; 89US-0323777.

XX (BIOM-) BIOMEASURE INC.

XX Eck CR, Moreau S;

XX WPI; 1990-292105/39.

PT New octa:peptide somatostatin analogues - useful for treating
 PT diabetes, liver disease, growth hormone dependent cancer,
 PT alzheimers disease etc.

PS Claim 5; Page 5; 8pp; English.

CC Analogues to somatostatin, useful for reducing or inhibiting release
 CC of antagonising growth hormone, insulin, glucagon or pancreatic
 CC exocrine solution. The peptides can treat diabetes related retinopathy,
 CC cirrhosis, hepatitis, Alzheimers disease, cancers etc.

XX Sequence 8 AA;

Query Match 92.6%; Score 50; DB 11; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CYMKVCT 8
 |||||
 DB 2 CYMKVCT 8

RESULT 13

AA07021
 ID AAR07021 standard; protein; 8 AA.

AC AAR07021;

DT 21-JAN-1991 (first entry)

DE Octapeptide somatostatin analogue.

KW Growth hormone dependent cancer; diabetes; Alzheimers disease;
 KM Cirrhosis; hepatitis; insulin; glucagon.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /label= OTHER, Dextrorotatory
 FT /note= "D-Beta-Nal"

FT Modified-site 3 /label= Iodinated phenyl ring

FT Modified-site 4 /label= Dextrorotatory

XX EP389180-A.

XX 26-SEP-1990.

XX 15-MAR-1990; 90EP-0302760.

XX 15-MAR-1989; 89US-0323777.

XX (BIOM-) BIOMEASURE INC.

XX Eck CR, Moreau S;

XX WPI; 1990-292105/39.

PT New octa:peptide somatostatin analogues - useful for treating
 PT diabetes, liver disease, growth hormone dependent cancer,
 PT alzheimers disease etc.

PS Claim 5; Page 5; 8pp; English.

CC Analogues to somatostatin, useful for reducing or inhibiting release
 CC of antagonising growth hormone, insulin, glucagon or pancreatic
 CC exocrine solution. The peptides can treat diabetes related retinopathy,
 CC cirrhosis, hepatitis, Alzheimers disease, cancers etc.

XX Sequence 8 AA;

Query Match 92.6%; Score 50; DB 11; Length 8;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CYMKVCT 8
 |||||
 DB 2 CYMKVCT 8

RESULT 14

AA07024
 ID AAR07024 standard; protein; 8 AA.

```

AC AAR07024;
XX
XX 21-JAN-1991 (first entry)
XX
XX Octapeptide somatostatin analogue.
XX
XX Growth hormone dependent cancer; diabetes; Alzheimers disease;
XX Cirrhosis; hepatitis; insulin; glucagon.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1
XX /label= OTHER, Dextrorotatory
XX /note= "N-Ac-D-beta-Nal"
XX
XX Modified-site 3
XX /label= Iodinated phenyl ring
XX Modified-site 4
XX /label= Dextrorotatory
XX
XX EP389180-A.
XX
XX 26-SEP-1990.
XX
XX 15-MAR-1990; 90EP-0302760.
XX
XX 15-MAR-1989; 89US-0323777.
XX
XX (BIOM-) BIOMEASURE INC.
XX
XX Eck CR, Moreau S;
XX
XX WPI; 1990-292105/39.
XX
XX New octe:peptide somatostatin analogues - useful for treating
XX diabetes, liver disease, growth hormone dependent cancer,
XX alzheimers disease etc.
XX
XX Claim 5; Page 5; 8pp; English.
XX
XX Analogues to somatostatin, useful for reducing or inhibiting release
XX of antagonising growth hormone, insulin, glucagon or pancreatic
XX exocrine solution. The peptides can treat diabetes related retinopathy,
XX cirrhosis, hepatitis, Alzheimer's disease, cancers etc.
XX
XX Sequence 8 AA;
SQ
Query Match 92.6%; Score 50; DB 11; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CYWKVCT 8
DB 2 CYWKVCT 8

```

```

RESULT 15
AAR10504
ID AAR10504 standard; Protein; 8 AA.
AC
XX
XX AAR10504;
XX
XX 25-MAR-2003 (updated)
DT 10-MAR-2003 (updated)
DT 09-APR-1991 (first entry)
XX
XX Pamoate of somatostatin octapeptide.
XX
XX Sustained release peptide compsn.; lactic acid; glycolic acid;
XX somatostatin; luteinising hormone releasing hormone.
XX
XX Unidentified.
XX

```

```

FH Key Location/Qualifiers
FT Modified-site 1..1
FT /label= Phe= D-Phe
FT Modified-site 4..4
FT /label= Trp= D-Trp
FT Disulfide-bond 2..7
XX
XX GB2234169-A.
XX
XX 30-JAN-1991.
XX
XX 17-JUL-1990; 90GB-0015722.
XX
XX 28-JUL-1989; 89CH-0002829.
XX
XX (DEBI-) DEBIOPHARM SA.
XX
XX Orsellini P;
XX
XX WPI; 1991-031530/05.
XX
XX Sustained release pharmaceutical peptide compsn. prepn. -
XX involves preheating and compressing mixt. of lactic and glycolic
XX acid copolymers with active substance pamoate, etc.
XX
XX Claim 9; page 12; 14pp; English.
XX
XX This peptide is used as the active substance in a pharmaceutical
XX compsn. for sustained release of somatostatin. It is mixed with a
XX lactic acid/glycolic acid co-polymer. All components are in the
XX form of microparticles and an injectable soln. is obtd. See also
XX AAR10501-03 and AAR10505-09.
XX (Updated on 10-MAR-2003 to add missing OS field.)
XX
XX Sequence 8 AA;
SQ
Query Match 92.6%; Score 50; DB 12; Length 8;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CYWKVCT 8
DB 2 CYWKVCT 8

```

```

Search completed: February 11, 2004, 06:31.48
Job time : 42 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 06:32:43 ; Search time 34 Seconds
(without alignments)
49.266 Million cell updates/sec

Title: SEQ1
Perfect score: 54
Sequence: 1 acywkvc 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : *

```
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	83.3	9	14	US-10-036-869-13
2	45	83.3	33	12	US-10-300-083-3
3	45	83.3	34	14	US-10-036-869-15
4	41	75.9	27	12	US-10-058-053A-263
5	41	75.9	78	12	US-10-058-053A-38
6	41	75.9	78	12	US-10-058-053A-212
7	40	74.1	8	12	US-10-239-701-11
8	40	74.1	10	10	US-09-852-870A-10
9	40	74.1	21	12	US-10-058-053A-234
10	40	74.1	40	12	US-10-058-053A-231
11	40	74.1	40	12	US-10-058-053A-264
12	40	74.1	45	12	US-10-058-053A-280
13	40	74.1	90	12	US-10-058-053A-80
14	40	74.1	93	12	US-10-058-053A-116

16	40	74.1	95	12	US-10-058-053A-89	Sequence 89, Appl
17	40	74.1	119	12	US-10-104-047-2041	Sequence 2041, Ap
18	40	74.1	5179	9	US-09-922-217-1068	Sequence 1068, Ap
19	40	74.1	5179	10	US-09-833-263-1068	Sequence 1068, Ap
20	40	74.1	5179	14	US-10-025-380-1068	Sequence 1068, Ap
21	38	70.4	8	10	US-09-758-426-3	Sequence 3, Appl1
22	38	70.4	8	10	US-09-758-426-3	Sequence 3, Appl1
23	38	70.4	8	10	US-09-758-198-3	Sequence 3, Appl1
24	38	70.4	8	11	US-09-861-661-3	Sequence 3, Appl1
25	37	68.5	7	10	US-09-758-128-1	Sequence 1, Appl1
26	37	68.5	7	10	US-09-758-426-1	Sequence 1, Appl1
27	37	68.5	7	10	US-09-758-198-1	Sequence 1, Appl1
28	37	68.5	7	11	US-09-861-661-1	Sequence 1, Appl1
29	37	68.5	8	15	US-10-079-384-35	Sequence 35, Appl
30	37	68.5	9	10	US-09-758-128-2	Sequence 2, Appl1
31	37	68.5	9	10	US-09-758-426-2	Sequence 2, Appl1
32	37	68.5	9	10	US-09-758-198-2	Sequence 2, Appl1
33	37	68.5	9	11	US-09-861-661-2	Sequence 2, Appl1
34	37	68.5	12	12	US-10-360-101-98	Sequence 98, Appl
35	37	68.5	21	12	US-10-058-053A-290	Sequence 290, App
36	37	68.5	21	12	US-10-058-053A-292	Sequence 292, App
37	37	68.5	21	12	US-10-058-053A-340	Sequence 340, App
38	37	68.5	22	12	US-10-058-053A-258	Sequence 258, App
39	37	68.5	22	12	US-10-058-053A-299	Sequence 299, App
40	37	68.5	23	9	US-09-864-761-38459	Sequence 38459, A
41	37	68.5	23	12	US-10-058-053A-244	Sequence 244, App
42	37	68.5	25	12	US-10-058-053A-245	Sequence 245, App
43	37	68.5	25	12	US-10-058-053A-252	Sequence 252, App
44	37	68.5	25	12	US-10-058-053A-259	Sequence 259, App
45	37	68.5	25	12	US-10-058-053A-260	Sequence 260, App

ALIGNMENTS

```
RESULT 1
US-10-036-869-13
; Sequence 13, Application US/10036869
; Publication No. US2002015156A1
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
; ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/036,869
; FILING DATE: 29-NOV-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-036-869-13

Query Match 83.3%; Score 45; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 7, 1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYWKVC 7
| | | | |
DB 3 CYWKVC 8

RESULT 2
US-10-300-083-3
Sequence 3, Application US/10300083
Publication No. US20030153502A1

GENERAL INFORMATION:

APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA
TITLE OF INVENTION: SYNTHETIC APPROACH TO DESIGNED CHEMICAL
STRUCTURES

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: MUEITING, RAASCH, GERHARDT & SCHWAPPACH, P.A.
STREET: 119 No. US20030153502A1h Fourth Street, Suite 203

CITY: Minneapolis

STATE: Minnesota

COUNTRY: U.S.A.

ZIP: 55401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/300,083

FILING DATE: 20-No. US20030153502A1-2002

*CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/194,296

FILING DATE: 15-Oct-1999

APPLICATION NUMBER: US 08/653,632

FILING DATE: 24-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: MCCORMACK, MYRA M.

REGISTRATION NUMBER: 36,602

REFERENCE/DOCKET NUMBER: 110,00330220

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-305-1225

TELEFAX: 612-305-1228

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-300-083-3

Query Match 83.3%; Score 45; DB 12; Length 33;
Best Local Similarity 100.0%; Pred. No. 4,3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYWKVC 7
| | | | |
DB 13 CYWKVC 18

RESULT 3
US-10-036-869-15

Sequence 15, Application US/10036869
Publication No. US20020151516A1

GENERAL INFORMATION:

APPLICANT: Mixson, James A

TITLE OF INVENTION: CARRIER: DNA COMPLEXES CONTAINING DNA

ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE

THERAPY

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Connolly, Bove, Lodge, & Hutz

STREET: 1220 Market Street, P.O. Box 2207

CITY: Wilmington

STATE: Delaware

COUNTRY: U.S.A.

ZIP: 19899

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/036,869

FILING DATE: 29-No. US20020151516A1-2001

*CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/985,526

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/608,845

FILING DATE: 16-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: McMorow Jr., Robert G

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302) 658-9141

TELEFAX: (302) 658-5613

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Query Match 83.3%; Score 45; DB 14; Length 34;
Best Local Similarity 100.0%; Pred. No. 4,4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYWKVC 7
| | | | |
DB 3 CYWKVC 8

RESULT 4
US-10-058-053A-263
Sequence 263, Application US/10058053A
Publication No. US20030170222A1

GENERAL INFORMATION:

APPLICANT: University of Utah Research Foundation

APPLICANT: Cognetix, Inc.

APPLICANT: Jones, Robert M.

APPLICANT: Garrett, James E.

APPLICANT: Watkins, Maren

APPLICANT: Oliveira, Baldomero M.

TITLE OF INVENTION: B-Superfamily Conotoxins

FILE REFERENCE: 2314-248

CURRENT APPLICATION NUMBER: US/10/058,053A

PRIOR FILING DATE: 2002-07-05

PRIOR APPLICATION NUMBER: US 60/264323

NUMBER OF SEQ ID NOS: 340

SOFTWARE: Patentin version 3.0

SEQ ID NO 263

LENGTH: 27

TYPE: PRT

ORGANISM: Conus sulcatus
US-10-058-053A-263

Query Match 75.9%; Score 41; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWKVC 7
Db 12 CYWKAC 17

RESULT 5
US-10-058-053A-38

Sequence 38, Application US/10058053A
Publication No. US20030170222A1
GENERAL INFORMATION:
APPLICANT: University of Utah Research Foundation
APPLICANT: Cognetix, Inc.
APPLICANT: Jones, Robert M.
APPLICANT: Garrett, James E.
APPLICANT: Watkins, Maren
APPLICANT: Oliveira, Baldomero M.
TITLE OF INVENTION: B-Superfamily Conotoxins
FILE REFERENCE: 2314-248
CURRENT APPLICATION NUMBER: US/10/058,053A
CURRENT FILING DATE: 2002-07-05
PRIOR APPLICATION NUMBER: US 60/264323
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 340
SOFTWARE: PatentIn version 3.0
SEQ ID NO 38
LENGTH: 78
TYPE: PRT
ORGANISM: Conus sulcatus
US-10-058-053A-38

Query Match 75.9%; Score 41; DB 12; Length 78;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWKVC 7
Db 62 CYWKAC 67

RESULT 6
US-10-058-053A-212

Sequence 212, Application US/10058053A
Publication No. US20030170222A1
GENERAL INFORMATION:
APPLICANT: University of Utah Research Foundation
APPLICANT: Cognetix, Inc.
APPLICANT: Jones, Robert M.
APPLICANT: Garrett, James E.
APPLICANT: Watkins, Maren
APPLICANT: Oliveira, Baldomero M.
TITLE OF INVENTION: B-Superfamily Conotoxins
FILE REFERENCE: 2314-248
CURRENT APPLICATION NUMBER: US/10/058,053A
CURRENT FILING DATE: 2002-07-05
PRIOR APPLICATION NUMBER: US 60/264323
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 340
SOFTWARE: PatentIn version 3.0
SEQ ID NO 212
LENGTH: 78
TYPE: PRT
ORGANISM: Conus sulcatus
US-10-058-053A-212

Query Match 75.9%; Score 41; DB 12; Length 78;
Best Local Similarity 83.3%; Pred. No. 33;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWKVC 7
Db 62 CYWKAC 67

RESULT 7
US-10-239-701-11

Sequence 11, Application US/10239701
Publication No. US20030153078A1
GENERAL INFORMATION:
APPLICANT: DHANAK, Dashedant
APPLICANT: KNIGHT, Steven D.
APPLICANT: WARREN, Gregory L.
TITLE OF INVENTION: Uroctension II Analogs
FILE REFERENCE: P51081
CURRENT APPLICATION NUMBER: US/10/239,701
CURRENT FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: PCT/US00/32408
PRIOR FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 60/168,141
PRIOR FILING DATE: 1999-11-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 8
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: cyclic disulfide
US-10-239-701-11

Query Match 74.1%; Score 40; DB 12; Length 8;
Best Local Similarity 71.4%; Pred. No. 7; 1e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACYWKVC 7
Db 1 ACYWKAC 7

RESULT 8
US-09-852-870A-10

Sequence 10, Application US/09852870A
Patent No. US20020165132A1
GENERAL INFORMATION:
APPLICANT: Goodman, Murray
APPLICANT: Osapay, George
TITLE OF INVENTION: Lanthionin Bridged Proteins
FILE REFERENCE: LKR 9122-D
CURRENT APPLICATION NUMBER: US/09/852,870A
CURRENT FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: US 09/384,601
PRIOR FILING DATE: 1999-08-26
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cys-Ser Lanthionine bridge
NAME/KEY: SITE
LOCATION: (1)..(8)
OTHER INFORMATION: Xaa= D-Phe, D-beta-Nal, or Phe
US-09-852-870A-10

Query Match 74.1%; Score 40; DB 10; Length 10;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACYWKVCT 8

Db 3 ACFWKVST 10

RESULT 9
US-09-852-870A-11
; Sequence 11, Application US/09852870A
; Patent No. US20020165132A1
; GENERAL INFORMATION:
; APPLICANT: Goodman, Murray
; APPLICANT: Osabay, George
; TITLE OF INVENTION: Lanthionin Bridged Proteins
; FILE REFERENCE: LKR 9122-D
; CURRENT APPLICATION NUMBER: US/09/852,870A
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/384,601
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cys-Ser Lanthionine bridge
; NAME/KEY: SITE
; LOCATION: (1)-(8)
; OTHER INFORMATION: Xaa- D-Phe, D-beta-Nal, or Phe
US-09-852-870A-11

Query Match 74.1%; Score 40; DB 10; Length 10;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACFWKVCT 8
Db 3 ACFWKVST 10

RESULT 10
US-10-058-053A-234
; Sequence 234, Application US/10058053A
; Publication No. US20030170222A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Oliveira, Baldomero M.
; TITLE OF INVENTION: B-Superfamily Conotoxins
; FILE REFERENCE: 2314-248
; CURRENT APPLICATION NUMBER: US/10/058,053A
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US 60/264323
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 234
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Conus geographus
US-10-058-053A-234

Query Match 74.1%; Score 40; DB 12; Length 21;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
Db 3 CFWKVST 9

RESULT 11
US-10-058-053A-231
; Sequence 231, Application US/10058053A
; Publication No. US20030170222A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Oliveira, Baldomero M.
; TITLE OF INVENTION: B-Superfamily Conotoxins
; FILE REFERENCE: 2314-248
; CURRENT APPLICATION NUMBER: US/10/058,053A
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US 60/264323
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 231
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Conus geographus
US-10-058-053A-231

Query Match 74.1%; Score 40; DB 12; Length 40;
Best Local Similarity 71.4%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
Db 22 CFWKVST 28

RESULT 12
US-10-058-053A-264
; Sequence 264, Application US/10058053A
; Publication No. US20030170222A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Oliveira, Baldomero M.
; TITLE OF INVENTION: B-Superfamily Conotoxins
; FILE REFERENCE: 2314-248
; CURRENT APPLICATION NUMBER: US/10/058,053A
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US 60/264323
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 264
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Conus docki
US-10-058-053A-264

Query Match 74.1%; Score 40; DB 12; Length 40;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWKVCT 7
Db 12 CYWVCT 17

RESULT 13
US-10-058-053A-280
; Sequence 280, Application US/10058053A
; Publication No. US20030170222A1
; GENERAL INFORMATION:

```

; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Olivera, Baldomero M.
; TITLE OF INVENTION: B-Superfamily Conotoxins
; FILE REFERENCE: 2314-248
; CURRENT APPLICATION NUMBER: US/10/058,053A
; PRIOR FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US 60/264323
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 280
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Conus cinereus
US-10-058-053A-280

```

```

Query Match          74.1%; Score 40; DB 12; Length 45;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 CYWKVCT 8
        |:|||
DB      12 CFWKST 18

```

```

RESULT 14
US-10-058-053A-80
; Sequence 80, Application US/10058053A
; Publication No. US20030170222A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Olivera, Baldomero M.
; TITLE OF INVENTION: B-Superfamily Conotoxins
; FILE REFERENCE: 2314-248
; CURRENT APPLICATION NUMBER: US/10/058,053A
; PRIOR FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US 60/264323
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 80
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Conus bocki
US-10-058-053A-80

```

```

Query Match          74.1%; Score 40; DB 12; Length 90;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 CYWKVC 7
        |:|||
DB      62 CYNVVC 67

```

```

RESULT 15
US-10-058-053A-116
; Sequence 116, Application US/10058053A
; Publication No. US20030170222A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren

```

```

; APPLICANT: Olivera, Baldomero M.
; TITLE OF INVENTION: B-Superfamily Conotoxins
; FILE REFERENCE: 2314-248
; CURRENT APPLICATION NUMBER: US/10/058,053A
; PRIOR FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US 60/264323
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 116
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Conus geographus
US-10-058-053A-116

```

```

Query Match          74.1%; Score 40; DB 12; Length 93;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 CYWKVCT 8
        |:|||
DB      72 CFWKST 78

```

```

Search completed: February 11, 2004, 06:38:00
Job time : 34 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 06:30:18 ; Search time 22 seconds
(without alignments)
15.386 Million cell updates/sec

Title: SEQ1

Perfect score: 54

Sequence: 1 acykwvct 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	92.6	8	1	US-08-286-748B-7
2	50	92.6	8	3	US-08-586-670A-13
3	50	92.6	8	3	US-08-586-670A-14
4	50	92.6	9	3	US-08-586-670A-11
5	50	92.6	9	3	US-08-586-670A-15
6	46	85.2	8	3	US-09-484-318-1
7	46	85.2	8	3	US-09-484-318-8
8	46	85.2	8	3	US-09-484-319-1
9	46	85.2	8	3	US-09-484-319-8
10	46	85.2	8	3	US-09-484-320-1
11	46	85.2	8	3	US-09-484-320-8
12	46	85.2	8	3	US-09-484-321-1
13	46	85.2	8	3	US-09-484-321-8
14	46	85.2	8	3	US-09-484-323-1
15	46	85.2	8	3	US-09-484-323-8
16	46	85.2	8	3	US-09-325-769-1
17	46	85.2	8	3	US-09-325-769-2
18	46	85.2	8	3	US-09-636-170-1
19	46	85.2	8	3	US-09-636-170-8
20	46	85.2	8	3	US-09-637-518-1
21	46	85.2	8	3	US-09-637-518-8
22	45	83.3	6	2	US-08-282-980B-8
23	45	83.3	6	3	US-08-931-095-5
24	45	83.3	9	3	US-08-985-526-13
25	45	83.3	3	2	US-08-671-487A-3
26	45	83.3	3	4	US-09-194-296-3
27	45	83.3	3	3	US-08-985-526-15

28	42	77.8	9	3	US-08-586-670A-16	Sequence 16, App1
29	41	75.9	8	4	US-08-586-670A-19	Sequence 19, App1
30	38	70.4	386	3	US-09-252-991A-20990	Sequence 20990, A
31	37	68.5	6	1	US-07-902-935-2	Sequence 2, App1
32	37	68.5	6	2	US-08-465-764-3	Sequence 3, App1
33	37	68.5	6	2	US-08-470-932-2	Sequence 2, App1
34	37	68.5	6	2	US-08-467-025-2	Sequence 2, App1
35	37	68.5	6	2	US-08-347-397-3	Sequence 3, App1
36	37	68.5	6	3	US-08-931-095-22	Sequence 22, App1
37	37	68.5	6	3	US-08-931-095-23	Sequence 23, App1
38	37	68.5	7	4	US-09-144-991B-38	Sequence 38, App1
39	37	68.5	7	4	US-09-144-991B-39	Sequence 39, App1
40	37	68.5	7	4	US-09-144-991B-40	Sequence 40, App1
41	37	68.5	7	4	US-09-144-991B-41	Sequence 41, App1
42	37	68.5	8	1	US-08-233-558-2	Sequence 2, App1
43	37	68.5	8	1	US-08-233-558-5	Sequence 5, App1
44	37	68.5	8	1	US-08-233-558-6	Sequence 6, App1
45	37	68.5	8	1	US-08-233-558-7	Sequence 7, App1

ALIGNMENTS

RESULT 1
US-08-286-748B-7
; Sequence 7, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gutewich
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 502 or 558X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Faase
; REGISTRATION NUMBER: 32,983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: Xaa is D- Naphthylalanine
US-08-286-748B-7
Query Match 92.6%; Score 50; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 2 CYWKVCT 8
DB 2 CYWKVCT 8

RESULT 2

US-08-586-670A-13
; Sequence 13, Application US/08586670A
; Patent No. 6241965
; GENERAL INFORMATION:
; APPLICANT: McBride, William
; APPLICANT: Dean, Richard T.
; TITLE OF INVENTION: Somatostatin Derivatives
; TITLE OF INVENTION: And their Radiolabeled Products
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,670A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6241965nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,385-DD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1234
; TELEFAX: 312-715-1000
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..4
; OTHER INFORMATION: /label= Variant residues
; OTHER INFORMATION: /note= "Xaa is D-naphthylalanine and is linked
; OTHER INFORMATION: to DTPA; Trp is in the D conformation;
US-08-586-670A-13
Query Match 92.6%; Score 50; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
DB 2 CYWKVCT 8

RESULT 3

US-08-586-670A-14
; Sequence 14, Application US/08586670A
; Patent No. 6241965
; GENERAL INFORMATION:
; APPLICANT: McBride, William
; APPLICANT: Dean, Richard T.
; TITLE OF INVENTION: Somatostatin Derivatives
; TITLE OF INVENTION: And their Radiolabeled Products
; NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,670A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6241965nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,385-DD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..4
; OTHER INFORMATION: /label= Variant residues
; OTHER INFORMATION: /note= "Xaa is D-naphthylalanine and is linked
; OTHER INFORMATION: to 2-ketoglutonyl; Trp is in the D conformation;
US-08-586-670A-14
Query Match 92.6%; Score 50; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
DB 2 CYWKVCT 8

RESULT 4

US-08-586-670A-11
; Sequence 11, Application US/08586670A
; Patent No. 6241965
; GENERAL INFORMATION:
; APPLICANT: McBride, William
; APPLICANT: Dean, Richard T.
; TITLE OF INVENTION: Somatostatin Derivatives
; TITLE OF INVENTION: And their Radiolabeled Products
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,670A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: No. 6241965nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 92,385-DD
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..3
OTHER INFORMATION: /label= Variant residues
OTHER INFORMATION: /note= "The Lys is linked to a BMT chelator
OTHER INFORMATION: through the side chain nitrogen; Xaa is
OTHER INFORMATION: D-naphthylalanine; the Cys sulfur is methylated;
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5..8
OTHER INFORMATION: /label= Variant residues
OTHER INFORMATION: /note= "The Trp residue is in the D conformation;
OTHER INFORMATION: the Cys side chain sulfur is methylated;
US-08-586-670A-11

Query Match 92.6%; Score 50; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
DB 3 CYWKVCT 9

RESULT 5
US-08-586-670A-15
Sequence 15, Application US/08586670A
Patent No. 6241965
GENERAL INFORMATION:
APPLICANT: McBride, William
TITLE OF INVENTION: Somatostatin Derivatives
TITLE OF INVENTION: And their Radiolabeled Products
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/586,670A
FILING DATE: 22-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6241965nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 92,385-DD
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..3
OTHER INFORMATION: /label= Variant residues
OTHER INFORMATION: /note= "Lys is linked to a BMT chelator
OTHER INFORMATION: through the side chain nitrogen and to DTPA
OTHER INFORMATION: at the N-terminus; Xaa is D-naphthylalanine;
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3..8
OTHER INFORMATION: /label= Variant residues
OTHER INFORMATION: /note= "The Trp residue is in the D conformation;
OTHER INFORMATION: each of the Cys side chain sulfur atoms are
OTHER INFORMATION: methylated;
US-08-586-670A-15

Query Match 92.6%; Score 50; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
DB 3 CYWKVCT 9

RESULT 6
US-09-484-318-1
Sequence 1, Application US/09484318
Patent No. 6180085
GENERAL INFORMATION:
APPLICANT: Achilefu, Samuel
APPLICANT: Rajagopalan, Raghavan
APPLICANT: Dorehow, Richard B
APPLICANT: Bugaj, Joseph E.
TITLE OF INVENTION: NOVEL DYES
FILE REFERENCE: DNA STRING
CURRENT APPLICATION NUMBER: US/09/484,318
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Octreotide
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: This is D-phenylalanine
NAME/KEY: DISULFID
LOCATION: (2)..(7)
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: This is D-tryptophan
NAME/KEY: SITE
LOCATION: (8)
OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-318-1

Query Match 85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CYWKVCT 8
DB 2 CYWKVCT 8

```
RESULT 7
US-09-484-318-8
; Sequence 8, Application US/09484318
; Patent No. 6180085
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorschow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: NOVEL DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,318
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-318-8
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
DB      2 CYWKCT 8
```

```
RESULT 8
US-09-484-319-1
; Sequence 1, Application US/09484319
; Patent No. 6180086
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorschow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: NEW HYDROPHILIC CYANINE DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,319
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotide
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-319-1
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
DB      2 CYWKCT 8
```

```
RESULT 9
US-09-484-319-8
; Sequence 8, Application US/09484319
; Patent No. 6180086
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorschow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: NEW HYDROPHILIC CYANINE DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,319
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-319-8
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
DB      2 CYWKCT 8
```

```
RESULT 10
US-09-484-320-1
; Sequence 1, Application US/09484320
; Patent No. 6180087
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorschow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: Tunable Indocyanine Dyes For Biomedical Applications
; FILE REFERENCE: dnastng
; CURRENT APPLICATION NUMBER: US/09/484,320
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotide
; NAME/KEY: SITE
; LOCATION: (1)
```

```
OTHER INFORMATION: This is D-phenylalanine
NAME/KEY: DISULFID
LOCATION: (2)..(7)
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: This is D-tryptophan
NAME/KEY: SITE
LOCATION: (8)
OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-320-1
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
        |||||
Db      2 CYWKCT 8
```

```
RESULT 11
US-09-484-320-8
```

```
Sequence 8, Application US/09484320
Patent No. 6180087
GENERAL INFORMATION:
APPLICANT: Achilefu, Samuel
APPLICANT: Rajagopalan, Raghavan
APPLICANT: Dorehow, Richard B
APPLICANT: Bugaj, Joseph E.
TITLE OF INVENTION: Tunable Indocyanine Dyes For Biomedical Applications
FILE REFERENCE: dnastring
CURRENT APPLICATION NUMBER: US/09/484,320
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Octreotate
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: This is D-phenylalanine
NAME/KEY: DISULFID
LOCATION: (2)..(7)
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: This is D-tryptophan
US-09-484-320-8
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
        |||||
Db      2 CYWKCT 8
```

```
RESULT 12
US-09-484-321-1
```

```
Sequence 1, Application US/09484321
Patent No. 6183726
GENERAL INFORMATION:
APPLICANT: Achilefu, Samuel
APPLICANT: Rajagopalan, Raghavan
APPLICANT: Dorehow, Richard B
APPLICANT: Bugaj, Joseph E.
TITLE OF INVENTION: VERSATILE HYDROPHILIC DYES
FILE REFERENCE: DNA STRING
CURRENT APPLICATION NUMBER: US/09/484,321
CURRENT FILING DATE: 2000-01-18
```

```
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Octreotide
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: This is D-phenylalanine
NAME/KEY: DISULFID
LOCATION: (2)..(7)
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: This is D-tryptophan
NAME/KEY: SITE
LOCATION: (8)
OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-321-1
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
        |||||
Db      2 CYWKCT 8
```

```
RESULT 13
US-09-484-321-8
```

```
Sequence 8, Application US/09484321
Patent No. 6183726
GENERAL INFORMATION:
APPLICANT: Achilefu, Samuel
APPLICANT: Rajagopalan, Raghavan
APPLICANT: Dorehow, Richard B
APPLICANT: Bugaj, Joseph E.
TITLE OF INVENTION: VERSATILE HYDROPHILIC DYES
FILE REFERENCE: DNA STRING
CURRENT APPLICATION NUMBER: US/09/484,321
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Octreotate
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: This is D-phenylalanine
NAME/KEY: DISULFID
LOCATION: (2)..(7)
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: This is D-tryptophan
US-09-484-321-8
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
        |||||
Db      2 CYWKCT 8
```

```
RESULT 14
```

```
US-09-484-323-1
Sequence 1, Application US/09484323
```

```
/ Patent No. 6190641
/ GENERAL INFORMATION:
/ APPLICANT: Achilefu, Samuel
/ APPLICANT: Rajagopalan, Raghavan
/ APPLICANT: Dorschow, Richard B
/ APPLICANT: Bugaj, Joseph E.
/ TITLE OF INVENTION: NOVEL INDOCYANINE DYES
/ FILE REFERENCE: DNA STRING
/ CURRENT APPLICATION NUMBER: US/09/484,323
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 8
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Octreotide
/ NAME/KEY: SITE
/ LOCATION: (1)
/ OTHER INFORMATION: This is D-phenylalanine
/ NAME/KEY: DISULFID
/ LOCATION: (2)..(7)
/ NAME/KEY: SITE
/ LOCATION: (4)
/ OTHER INFORMATION: This is D-tryptophan
/ NAME/KEY: SITE
/ LOCATION: (6)
/ OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-323-1
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 CYWKVCT 8
        |||||
Db       2 CYWKCT 8
```

```
RESULT 15
US-09-484-323-8
/ Sequence 8, Application US/09484323
/ Patent No. 6190641
/ GENERAL INFORMATION:
/ APPLICANT: Achilefu, Samuel
/ APPLICANT: Rajagopalan, Raghavan
/ APPLICANT: Dorschow, Richard B
/ APPLICANT: Bugaj, Joseph E.
/ TITLE OF INVENTION: NOVEL INDOCYANINE DYES
/ FILE REFERENCE: DNA STRING
/ CURRENT APPLICATION NUMBER: US/09/484,323
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8
/ LENGTH: 8
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Octreotide
/ NAME/KEY: SITE
/ LOCATION: (1)
/ OTHER INFORMATION: This is D-phenylalanine
/ NAME/KEY: DISULFID
/ LOCATION: (2)..(7)
/ NAME/KEY: SITE
/ LOCATION: (4)
/ OTHER INFORMATION: This is D-tryptophan
US-09-484-323-8
```

```
Query Match      85.2%; Score 46; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.5e+05;
```

```
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      2 CYWKVCT 8
        |||||
Db       2 CYWKCT 8
```

```
Search completed: February 11, 2004, 06:33:33
Job time : 22 secs
```